

Expressive Language in Preschoolers Born Preterm: Results of Language Sample Analysis and
Standardized Assessment

By

Caitlin Imgrund

Submitted to the graduate degree program in Speech-Language-Hearing: Sciences and Disorders
and the Graduate Faculty of the University of Kansas in partial fulfillment of the requirements
for the degree of Doctor of Philosophy.

Chair: Dr. Diane Loeb

Co-Chair: Dr. Holly Storkel

Dr. Debora Daniels

Dr. Nancy Brady

Dr. Gregory Cheatham

Date Defended: July 13, 2017

The dissertation committee for Caitlin McCormick Imgrund certifies that
this is the approved version of the following dissertation:

Expressive Language in Preschoolers Born Preterm: Results of Language Sample Analysis and Standardized Assessment

Chair: Diane Loeb

Date Approved: July 13, 2017

Abstract

Children born preterm constitute one of the largest populations of children at risk for the development of language impairments. A little over one in ten pregnancies result in a preterm birth and approximately 25% of these children go on to experience subsequent difficulties with language (CDC, 2015; Foster-Cohen, Friesen, Champion, & Woodward, 2010). Despite the high risk for language deficits in this population, few studies have investigated the conversational language skills of these children. In particular, the objective of this study was to investigate the grammatical and semantic skills of children born preterm via language sample analysis. A second aim of the study was to determine the relationship between conversational language skills and the results of standardized assessment of language in this population and investigate the role that non-linguistic factors such as attention and non-verbal intelligence play in standardized assessment results. Twenty-nine preschoolers born preterm and a comparison group of 29 full term peers participated in this study. The children in the preterm group performed more poorly than the full term group on measures of conversational semantic and grammatical skills obtained from language sample analysis. In contrast, the two groups performed similarly on all but one of the measures obtained from standardized assessments. The clinical implications of these findings are discussed.

Acknowledgements

First, I would like to thank the chair of my dissertation committee and my advisor, Dr. Diane Loeb. I started working with Dr. Loeb while I was still an undergraduate student and she encouraged me to pursue my PhD. She has challenged me to grow as a researcher, teacher and clinician. Her expertise in the area of language development and language disorders has been invaluable to my learning. I have gained so much from her guidance and I would like to thank her sincerely for the time and care she has invested in me.

Appreciation also goes to Dr. Holly Storkel, who graciously accepted the position of co-chair on my dissertation committee and allowed me to use her research facilities to store my data and conduct data analyses. She has been a tremendous resource during my time as a PhD student. My dissertation has also benefitted from the guidance of a perceptive and supportive committee. Thanks to the remainder of my committee: Dr. Debora Daniels, Dr. Nancy Brady, and Dr. Gregory Cheatham. This project was strengthened by their input and guidance.

For their help with data collection, transcription, and entry I would like to thank Cintamy Budiardjo, Veronica Fox, Abby Hill, and Callie Flading. For her consultation on Excel formulas and functions, thank you to Gaby Moran. For statistics guidance, thank you to Rebecca Swinburne of the Lifespan Institute.

I often joke with my friends and family that my parents have a predilection for discussing science at the dinner table. When I was growing up, acceptable dinner table topics included the molecular structure of table salt and the principles of heat convection. Although I never developed a passion for the so-called “hard” sciences, I would like to thank my mom (Jo) and my dad (Sam) for instilling in me an enthusiasm for the scientific method. They have also loved and

supported me unconditionally and for that, I am forever grateful. I have been lucky to have their encouragement.

Thank you to my Girl Scout leader, second mom, and SLP role model Kathy Johnston. Kathy introduced me to the field of speech-language pathology and encouraged me to become an SLP before I had even learned how to drive.

My husband, Bret, has been a constant source of support, companionship, and encouragement. I can always count on him for a big hug after a long day in the lab. He is a wonderful partner and I am so grateful to have him in my life. He has supported me in countless ways along this journey, especially in these last few months while I was on the job market and writing my dissertation.

My friends have been my ever-present cheering squad, and never let me doubt myself. Thank you to Katie Johnston, Jenny Smith, Serene Hakim Saliba, and Gaby Moran. I am so grateful for our friendship. My fellow PhD students have also offered encouragement, inspiration, and a good laugh when needed. Special thanks to Dr. Austin Rosner, Dr. Breanna Krueger, and Dr. Stephanie Knollhoff.

The completion of my dissertation was made possible through the financial support of several entities. I would like to acknowledge the financial support of the Council of Academic Programs in Communication Sciences and Disorders and the ASHA Foundation. Gratitude also goes to Dr. Steven Barlow for allowing me to recruit from his database of children born preterm.

Lastly, I would like to thank the children and families who participated in my dissertation study. Preterm birth brings many challenges, but also many triumphs. Thank you to these families and their wonderful children.

TABLE OF CONTENTS

INTRODUCTION	1
Neurodevelopmental Outcomes of Children Born Preterm.....	1
Cognitive Outcomes of Children Born Preterm.....	1
Fine and Gross Motor Outcomes in Children Born Preterm.....	2
Attention Outcomes in Children Born Preterm.....	2
Factors that Influence Neurodevelopmental Outcomes of Children Born Preterm.....	3
Language Outcomes of Children Born Preterm.....	4
Semantics.....	5
Morphology and Syntax.....	6
Pragmatics.....	8
Risk and Rates of Language Impairments during the Preschool Period.....	9
Assessment of Language in the Preschool Period.....	10
Language Sample Analysis.....	10
Standardized Language Assessment.....	17
Associations between Language Sample Analysis and Standardized, Norm- Referenced Assessment.....	19
Language Sample Analysis and Standardized Assessment in Children Born Preterm....	23
Rationale for Current Study.....	25
Research Questions.....	27
METHODS.....	29

Participants.....	29
Procedures.....	31
Evaluation of Language: Language Sample Analysis.....	32
Transcription and Coding Reliability Analyses.....	36
Evaluation of Language: Formal Standardized Assessment.....	38
Attention.....	39
Other Assessment Measures.....	41
RESULTS.....	41
Statistical Analyses.....	41
Analysis of Variance (ANOVAs).....	44
Research Question 1.....	44
Standardized Assessment of Language Results.....	44
Language Sample Analysis Results.....	45
Research Question 2.....	46
Attention/Hyperactivity Results.....	46
Research Question 3.....	47
Impact of Non-Linguistic Factors.....	47
DISCUSSION.....	49
Clinical Implications.....	55
Limitations of the Current Study.....	57
Future Directions.....	58
REFERENCES.....	60
APPENDIX A.....	73

APPENDIX B.....	74
APPENDIX C.....	80

Introduction

Preterm birth, or birth that occurs before 37 weeks gestational age (GA), occurs in approximately 1 out of every 10 live births in the United States each year (CDC, 2015). Despite advances in the perinatal care of these children in neonatal intensive care units (NICUs), rates of morbidity related to preterm birth have remained relatively stable over time (Fanaroff et al., 2007, Stoll et al., 2015). Due to the introduction of medical interventions including the use of antenatal corticosteroids and postnatal surfactant to help children with immature lungs breathe, survival rates of children born preterm have dramatically increased since the late 1980s (Paranka, Yoder, & Brehm, 1999; Polin, Carlo, & Committee on Fetus and Newborn, 2014). Improvements in the survival rates of children born preterm in the presence of stable morbidity rates has resulted in an increase in the number of NICU survivors who are at risk for adverse neurodevelopmental sequelae. Indeed, children born preterm constitute one of the largest groups of children at risk for poor neurodevelopmental outcomes, including language outcomes. Recent research has been conducted to better understand the neurodevelopmental outcomes and trajectories of children from this at-risk population and will be presented in the following sections.

Neurodevelopmental Outcomes of Children Born Preterm

Cognitive outcomes of children born preterm. Although there are children born preterm who exhibit little or no measurable deficits as a result of their premature birth, neurodevelopmental outcomes are frequently negatively influenced by preterm birth. Even when sociodemographic risk is taken into consideration, children born preterm remain at a significantly higher risk for below-average intelligence quotients (IQs) (i.e., IQs below 85) (Hack, Klein, & Taylor, 1995). Kerr-Wilson, Mackay, Smith, and Pell (2012) conducted a meta-

analysis of 27 studies to determine the association between preterm birth and subsequent cognitive abilities. The findings from this meta-analysis indicated that children born preterm have IQs on average 11.94 points lower than their peers who were born full term. Gestational age at birth was found to play an important linear role in later cognitive ability, with average IQs falling consistently for every one week decrease in gestation. These cognitive deficits have been shown to impact learning, with as many as 65% of children born at or before 32 weeks gestational age identified with learning disabilities, as compared to 13% of children born full term (Grunau, Whitfield, & Davis, 2002). A longitudinal study following a large cohort of children who were born preterm from age 6 to age 12 found persistent deficits in the preterm group on measures of general cognitive functioning (Mangrin, Horwood, & Woodward, 2017). This study also used growth curve modeling to show an overall trend of relative stability in the IQ trajectories of the children in the preterm group, indicating that for many children born preterm, increases over the course of development to attain typical levels do not occur.

Fine and gross motor outcomes in children born preterm. Difficulties in both fine and gross motor development are also common in children who were born preterm. Between 6% and 9% of children born at or before 32 weeks gestational age have a diagnosis of cerebral palsy (CP), and the rate of CP increases to 16%-28% when the child is born at or before 26 weeks gestational age (Milligan, 2010). Even children born without gross neurological damage resulting in CP have increased difficulty with gross motor and fine motor skills and deficits in coordination when compared to children who were born full term (Bos, Van Braeckel, Hitzert, Tanis, & Roze, 2012; Schmidhauser, Caflisch, Rousson, Bucher, & Latal, 2006).

Attention outcomes in children born preterm. High rates of attention difficulties have been identified in children born preterm (Bhutta, Cleves, Casey, Cradock, & Anand, 2002).

Difficulty with attention has been documented early in infancy as well as later in life (Butcher, Kalverboer, Geuze, & Stemmelaar, 2002). A review of the literature conducted by van de Weijer-Bergsma, Wijnroks, and Jongmans (2008) highlighted attention-related deficits in toddlers and preschoolers born preterm, including less mature visual orienting, shorter periods of sustained attention, and deficits in executive control of attention. These deficits in the area of attention do not appear to end after the preschool period and may manifest themselves as attention deficit hyperactivity disorder (ADHD) later in life. A meta-analysis conducted by Bhuta et al. (2002) on the impact of preterm birth on attention-related deficits indicated that children who were born preterm were at an increased risk for ADHD as well as other behavioral problems.

Factors that Influence Neurodevelopmental Outcomes of Children Born Preterm

Children born preterm are a heterogeneous population, and not all children who were born preterm are at an equal risk for poor neurodevelopmental outcomes. Recent studies have highlighted many factors that are correlated with poor neurodevelopmental outcomes in this population. These factors include gestational age at birth and birthweight, medical comorbidities, and environmental factors.

Birthweight and gestational age at birth are frequently cited in the literature as gross predictors of neurodevelopment in children who were born preterm. These studies indicate a consistent relationship between birthweight, gestational age, and neurodevelopmental outcomes (Curry, Pfeiffer, Slopen, & McVeigh, 2012; Kerr-Wilson et al., 2012).

Although birth gestational age and birthweight are good gross predictors of neurodevelopmental outcomes, other factors such as medical comorbidities have also been found to influence outcomes, after controlling for the effects of gestational age and birthweight.

Children who experience brain damage during the perinatal period such as intraventricular hemorrhage (IVH) or periventricular leukomalacia (PVL), commonly have significant motor, cognitive, and/or communication deficits (de Vries, Eken, Groenendaal, Haastert, & Meiners, 1993; Ment, Allan, Makuch, & Vohr, 2005; Mukerji, Shah, & Shah, 2015). Respiratory problems including chronic lung disease (CLD) are another medical comorbidity commonly seen in the preterm population. CLD and its resulting diffuse white matter damage is a risk factor for language and other neurodevelopmental deficits (Lewis, Singer, Fulton, Salvator, Short, Klein, & Baley, 2002; Singer, Siegel, Lewis, Hawkins, Yamashita, & Baley, 2001; Wickremasinghe et al., 2012).

One of the most important factors that influences the development of a child is his or her socioeconomic environment (Hart & Risley, 1992; Walker, Greenwood, Hart, & Carta, 1994). This is also true for children who are born preterm. Indeed, the negative effects of preterm birth and the comorbidities commonly experienced in the NICU may be partially mitigated by a high-quality socioeconomic environment (Scheiner & Sexton, 1991; Wickremasinghe et al., 2012; Wild, Betancourt, Brodsky, & Hurt, 2013). Wild et al. (2013) found that 45% of toddlers who were born preterm from low-socioeconomic families had mild to moderate language delays in comparison to just 8% of children from higher-socioeconomic families matched for gestational age, birthweight, and medical risk.

Language Outcomes of Children Born Preterm

As with other areas of neurodevelopment, language deficits in children born preterm have been well-documented across the life-span (Casiro, Moddemann, Stanwick, Panikkar-Thiessen, Cowan, & Cheang, 1990; Foster-Cohen, Edgin, Champion, and Woodward, 2007; Magill-Evans, Harrison, Van der Zalm, & Holdgrafer, 2002; van Noort-van der Spek, Franken, & Weisglas-

Kupers, 2012). These studies indicate a consistent negative relationship between preterm birth and subsequent language outcomes as well as a heightened risk for language impairment. Because the focus of this study was on preschool-aged children, this review of the literature emphasizes research conducted with children born preterm who are in the preschool age-range (i.e., ages 3-5 years).

The preschool period is a time of rapid language growth as children become increasingly able to use language to make their wants and needs known as well as to express their thoughts and feelings (Pence Turnbull & Justice, 2011). Many important language-related milestones are achieved during the preschool period for children with typical language development (Brown, 1973). These achievements, which occur in the areas of syntax, morphology, phonology, semantics, and pragmatics, provide a strong foundation for future language development.

Semantics. During the preschool period, children are rapidly acquiring new vocabulary and building their lexicons. Preschool-aged children who were born preterm lag behind their full-term peers in terms of receptive and expressive vocabulary development (Carvale, Tozzi, Albino, & Vicari, 2005; Foster-Cohen et al., 2010; Kilbride, Thorstad, & Daily, 2004). Given that early semantic abilities can predict later achievement in reading comprehension and academic success, semantics is an important area of language development (Roth, Speece, & Cooper, 2002).

Kilbride, Thorstad, and Daily (2004) investigated neurodevelopmental outcomes, including language outcomes, in extremely low birthweight children (ELBW) (i.e., birthweight below 801 grams). Their sample was comprised of 25 children born ELBW and 25 of their full-term siblings. At 3 and/or 5 years of age, the children were administered the Peabody Picture Vocabulary Test (PPVT), an assessment of receptive vocabulary. The children who were born ELBW scored, on average, 10.2 standard score points below their full term siblings on the

PPVT. Carvale et al. (2005) also assessed receptive semantic skills in children born preterm using the PPVT. The sample in this study consisted of 30 Italian children born between 30 and 34 weeks gestational age with no major neurological disabilities and 30 full term controls matched for age, sex, and parental education level. The children were assessed when they were between 3 and 4 years of age. The children born preterm performed more poorly on the PPVT than their full term peers. This finding is of particular interest because the children in the preterm group were considered “low-risk” (i.e., had normal cerebral ultrasounds, no major neurosensory impairments, and were not extremely low birthweight).

Expressive semantic skills have also been assessed in children born preterm (Foster-Cohen et al., 2010). Foster-Cohen et al. (2010) assessed expressive vocabulary in preschoolers who were born before 33 weeks gestational age and a group of age-matched full term peers. In this study, the children born preterm had statistically significant lower scores on the Formulating Labels subtest of the CELF-P, an expressive vocabulary subtest. However, the effect size was modest (.31) and smaller than the effect sizes found on the other subtests of the CELF-P.

Morphology and syntax. For children with typical language development, many gains in the area of morphology and syntax can be seen as children begin using longer and more grammatically complex utterances (Brown, 1973). Of particular interest to those studying preschoolers at risk for language impairments is the production of finite verb morphology. Although many domains of language are negatively impacted in children with language impairment, difficulty with finite verb morphology, or verbs and verb inflections that carry information about tense and agreement, is considered a hallmark characteristic of language impairment (Leonard, Caseli, Bortolini, McGregor, & Sabbadini, 1992). With respect to this proposed study, syntax refers to the application of rules for combining words into sentences,

morphology refers to the application of rules for combining parts of words carrying meaning (i.e., morphemes) into longer words, and morphosyntax refers to the application of rules governing morphemes that carry information about tense and agreement.

A study conducted by Le Normand and Cohen (1999) evaluated the effects of preterm birth on the production of finite and nonfinite verb morphology. The children who participated in the study were French-speaking preschoolers between the ages of 3;6 and 5 years. Language samples were gathered from the children and their productions of main verbs, auxiliary verbs/copulas, and non-finite verbs (e.g., infinitives, past participles, etc.) were recorded. These researchers noted a nearly 18 month delay in the acquisition of auxiliaries and copulas in children born preterm when compared to full term controls during the preschool period. These differences were not found for non-finite verb morphology. Although morphosyntax is an important domain of language development, few studies have been conducted in this area of language in children who were born preterm. More studies need to be conducted to better understand how preterm birth may impact subsequent morphological and syntactic abilities.

Sansavini et al. (2010) investigated syntactic abilities in Italian-speaking children born preterm at three and a half years of age. The 70 children in the preterm group had a mean gestational age of 30 weeks and did not have any major cerebral damage. A group of 40 healthy, full term preschoolers matched for age and parental education level were recruited to serve as a control group. To evaluate syntax in this study, the children were administered an Italian standardized assessment of syntax called the Prova di Ripetizione di Frasi (PRF). The PRF is a sentence repetition task in which the child is asked to repeat verbatim a series of sentences of varying length containing targeted syntactic structures. The mean length of utterance (MLU) used by the children in the task also can be calculated. The children in the preterm group had

lower PRF MLUs than the children in the full term group (3.7 compared to 4.2). Their errors were characterized by omissions across word types (i.e., nouns, verbs, function words), with function words being the most commonly omitted word type. It should be noted that although the PRF is primarily an assessment of syntax, sentence repetition tasks such as the PRF also tax the child's short-term phonological system.

No studies designed to specifically evaluate morphology and or syntax in English-speaking preschoolers born preterm were identified in this review of the literature. However, one study evaluating the broad-based language skills of preschoolers born preterm reported scores from the subtests of the CELF-P, which contains a subtest assessing expressive morphology (i.e., Word Structure subtest) and a subtest assessing expressive syntax via sentence repetition (i.e., Recalling Sentences subtest). In this study, Foster-Cohen et al. (2010) found statistically significant lower scores on the Recalling Sentences subtest ($d = .42$) and the Word Structure subtest ($d=.33$) in comparison to a group of full term children matched for socioeconomic status; however, both results had small effect sizes

Pragmatics. As with other areas of language development, children are also rapidly developing their pragmatic language skills during the preschool period. Unfortunately, very few studies have been conducted in the area of pragmatics in children born preterm, despite studies indicating higher risk in children born preterm for disorders that strongly impact pragmatic ability (i.e., autism) (Limperopoulos et al., 2008). Guirini et al. (2016) conducted a study assessing two aspects of pragmatic language with preschoolers born preterm. This study, which included 60 children born very preterm and a full term comparison group, assessed the preschooler's comprehension of idioms and production of narratives. The children in the preterm group understood fewer idioms and produced less cohesive narratives (i.e., they included more

tangential or conceptually incongruent utterances in their retells). This study only investigated two relatively narrow measures of pragmatic language ability; more studies need to be conducted in this domain of language before wider conclusion can be drawn about the pragmatic language skills of preschoolers born preterm. Although there are few studies investigating the pragmatic language skills of preschool-aged children born preterm, one study (Reidy et al., 2013) assessed pragmatics in slightly older children (i.e., mean age of 7 years). The parents of the children in the preterm group were given a questionnaire containing items about their child's pragmatic language skills (i.e., the Pragmatics Profile from the CELF). This study found statistically significant differences in parent-reported pragmatic abilities between children born preterm and children born full term, with children born full term reported as having more advanced pragmatic abilities than their preterm counterparts.

Risk and rates for language impairments during the preschool period. Several studies have investigated the overall rates and risk of language impairments in preschoolers who were born preterm. Foster-Cohen et al. (2010) found a high prevalence of mild to moderate language impairment in four-year-old children who were born preterm. In their sample of 100 children born preterm, 16% had a mild language delay (1–1.5 SD below the mean on the CELF-P) and 15% had a moderate language delay (1.5–2 SD below the mean on the CELF-P) in comparison to 8.6% and 6.7% respectively in the full term comparison group. Sansavini et al. also investigated rates of language impairment in children born preterm but without significant cerebral damage at 3;6 years of age. Their sample had similar rates of language impairment, with approximately one third (34.4%) of the children in the preterm group demonstrating significant delays in language acquisition in comparison to 7.5% of children in the full term comparison group. Taken together, these studies provide evidence that there is a high prevalence of mild to

moderate language deficits in children who were born preterm. Although these two studies provide evidence that children born preterm may be at an increased risk for language impairment, these studies assessed language ability solely through standardized assessment. While the administration of standardized tests is an important component of language assessment, it is generally recommended that other methods, such as language sample analysis, be used in conjunction with standardized testing (Costanza-Smith, 2010; Paul & Norbury, 2012).

Assessment of Language in the Preschool Period

Because children who were born preterm are at an increased risk for language-related difficulties, accurate assessment of language skills is particularly crucial in this population in order to identify children for early intervention. Researchers and clinicians assessing the language abilities of young children have a wide variety of methods available to them. Two commonly used methods for evaluating the language skills of young children are language sample analysis and the administration of standardized language assessments. Each of these methods has their own benefits and drawbacks, and it is generally recommended that these two evaluation procedures be used in conjunction with one another during the assessment of a child's language skills (Costanza-Smith, 2010; Paul & Norbury, 2012).

Language sample analysis. The analysis of children's utterances in spontaneous speech has long been used as a tool in the study of child language (Brown, 1973). When gathering a language sample, the researcher has several questions and concerns to consider, including 1) how the language sample will be collected, and 2) how the language sample will be analyzed.

The way in which a language sample is collected plays an important role in the overall quality of the language sample and research has been conducted to determine how language sampling contexts and procedures affect the language sample. Southwood and Russell (2004)

compared three different language sampling contexts; free play, story generation, and conversation. They found that although the five-year-old children in their sample used more overall utterances during free play, they produced the most syntactically complex language and longest utterances in the story generation context. Similarly, Wagner, Nettelbladt, Sahlen, and Niholm (2000) found that in preschoolers, a narrative task elicited more grammatical morphemes per utterance than a conversational task. However, the children were more fluent and intelligible during the conversational task.

Costanza-Smith (2010) suggests that a variety of language contexts and materials be used during the elicitation of a language sample in order to obtain a more complete and representative picture of the child's linguistic abilities. For preschool-aged children who are still developing language skills and the ability to use decontextualized language, Owens (2010) recommends supplying some contextual support when gathering a language sample and using familiar activities, topics, and materials. The materials used during the elicitation of the language sample may also influence the child's performance. For example, O'Brien and Nagle (1987) found that play with dolls elicited more complex language from children than when playing with vehicles.

An additional factor for the researcher to consider is the length of the language sample. Most researchers and clinical experts suggest gathering between 50 and 100 complete and intelligible utterances (Heilmann, 2010; Miller, 1981; Paul & Norbury 2012). Gavin and Giles (1996) found that using 50 complete and intelligible utterances produced significant test-retest correlations, but a sample of 175 utterances was needed to reach test-retest coefficients greater than .90. Heilmann, Nockerts, & Miller (2010) analyzed language samples of three different lengths (i.e., 1 minute, 3 minutes, and 7 minutes). In this study, length of the sample did not yield significant difference between the language sample measures (e.g., number of different words,

MLU, words per minute); however, stability of language measures improved as sample time increased for the preschoolers in the sample

Finally, the researcher should consider who the child is interacting with during the language sample collection. Some researchers choose to collect language samples while the child is interacting with a familiar conversational partner such as a parent or a peer (Eisenberg & Guo, 2015; Demuth, 1984). Others opt to collect language samples while the child is interacting with an unfamiliar conversational partner, such as the researcher (Souto, Leonard, & Deevy, 2014; Rice & Wexler, 1996). Currently, there is no consensus on the effect of the conversational partner in the quality of the language sample.

After the language sample has been collected, the researcher must next consider how the language sample will be transcribed and what analyses to conduct. Miller and Chapman (2000) have proposed standardized conventions for the transcription of language samples, and these conventions are widely implemented in the transcription of language samples in both the research and clinical setting.

When appropriate language sampling and transcription techniques are utilized, the language sample provides a representative picture of the child's expressive language skills. However, the language sample is of little use to the researcher or clinician if it is not analyzed. Many options exist for the analysis of language samples to determine the functioning of the child's semantic, syntactic, and morphological language skills. Recently, computerized programs such as the Systematic Analysis of Language Transcripts (SALT)(Miller & Chapman, 2012) have provided greater efficiency to researchers and clinicians who wish to analyze language samples.

Analysis of semantics via language sample analysis. A researcher or clinician who is interested in a child's expressive semantic skills may use a number of different analyses to assess the child's semantic skills within a discourse context. Number of different word roots, total number of words, and type token ratio (i.e., number of different words divided by the total number of words) are frequently calculated in the assessment of a child's semantic skills (Malvern & Richards, 2012; Watkins, Kelly, Harebers, & Hollis, 1995). Particularly, the calculation of number of different words roots is considered a robust assessment of a child's lexical diversity, and has been found to better differentiate children with language impairment from children with typical language skills than other measures such as type token ratio (Watkins et al., 1995).

Analysis of morphology and syntax via language sample analysis. The calculation of mean length of utterance (MLU) is perhaps one of the most commonly used measures of a child's expressive syntax, and some research has been conducted to determine its validity as a measure of syntactic ability (Eisenberg, Fersko, & Lundgren, 2001; Klee, 1992; Leonard & Finneran, 2003). Although MLU is considered too global of a language measure to be solely used in the diagnosis of language impairment, MLU is a valuable measure for assessing children's productive language skills, especially when used in conjunction with other measures (Leonard & Finneran, 2003; Paul & Norbury, 2012). If a particular grammatical structure or structures are of interest to the clinician or researcher, percent correct use in obligatory contexts can also be easily calculated from the language sample. Because the assessment of finite verb morphology is particularly useful in the diagnosis of language impairment, percent correct use of these types of verbs and morphemes is frequently calculated (Leonard, 2014; Rice & Wexler, 1996).

Other, more complicated systems for quantifying a child's use of morphology and syntax in conversational speech also exist. Two of these systems are Developmental Sentence Score (DSS) (Lee, 1974) and Index of Productive Syntax (IPSyn)(Scarborough, 1990). DSS is a procedure for estimating a child's syntactic (and to a minimal extent, semantic) abilities and includes information on eight syntactic categories: indefinite pronouns/noun modifiers, personal pronouns, main verbs, secondary verbs, negatives, conjunctions, interrogative reversals, and wh-questions. For every utterance, the child is given a point value between 1 and 8 for each of the syntactic categories used in the utterance. Lower point values are assigned to early-acquired syntactic constructions and higher point values assigned for later-acquired syntactic constructions. For example, uninflected main verbs (e.g., see, kick, go) are awarded one point and more complex verbal constructions such as have + been + verb + ing (e.g., have been walking) are awarded eight points. In addition, each sentence is given a sentence point of 1 or 0 depending on its semantic and grammatical correctness compared to an adult standard. The average DSS for the sample can then be calculated in an attempt to quantify the child's expressive syntactic skills. IPSyn is similar to DSS in that it also attempts to quantify morphosyntactic skill through analysis of syntactic constructions used by the child. In IPSyn, a sample of 100 utterances is analyzed for the presence of 56 syntactic structures (e.g., negation, noun phrases, etc.). The child is awarded points for using each of the syntactic structures up to two times, with a maximum score of 112.

Advantages of language sample analysis. The use of language sample analysis in the assessment of children's language skills has many advantages. First and foremost, language sample analysis is considered an ecologically valid assessment tool. Ecological validity refers to the extent to which the findings of an assessment are relevant to real-life functioning. The

analysis of a child's spontaneous language allows researchers and clinicians alike to determine how the child is actually using language in natural contexts (Costanza-Smith, 2010).

Additionally, language sample analysis has been determined to be a valid assessment tool for children for whom standardized assessment may present problems. This includes children from diverse linguistic backgrounds and children with attention and behavior deficits (Redmond, 2003; Rojas & Iglesias, 2010). The gathering of language samples places very few performance demands upon the child, because the person gathering the language sample can follow the child's lead and record the child's utterances. This can be beneficial for children who have difficulty complying with the performance demands required by many standardized, norm-referenced assessments (e.g., sustained attention, sitting at a table, complying with instructions, etc.).

Lastly, and very importantly, there is some evidence that data obtained from language sample analysis is more accurate in the diagnosis of language impairment in children than the use of standardized assessments. For example, Aram, Morris, and Hall (1993) investigated the overlap between measures of language ability (both standardized assessments and language sample analysis measures) and clinical diagnoses of specific language impairment (SLI). The results of this study indicated that MLU obtained from a language sample was a more sensitive measure than the standardized assessment results for positively identifying children who had been clinically diagnosed with SLI. A follow-up analysis conducted by Dunn, Flax, Sliwinski, and Aram (1996) investigated the conversational language skills of young children clinically diagnosed with SLI, but whose standardized assessment scores were not low enough to meet eligibility guidelines for SLI. These children showed significant difficulty with language measures obtained through language sample analysis. Specifically, using a combination of MLU

and percentage of structural errors (e.g., grammatical errors) was more optimal for diagnosing language impairment than any of the standardized tests used. The researchers involved in this study concluded that language sample analysis measures are more closely aligned with clinician's perceptions of language impairment in preschoolers than results from standardized assessment.

Disadvantages of language sample analysis. Some potential pitfalls exist in the use of language sample analysis to assess the language skills of preschool-aged children. First, language sample analysis may not be a useful assessment tool if the researcher is interested in a particular grammatical structure or structures, some of which may be less frequent in conversational speech than others. This is particularly problematic if the researcher is interested in determining percent correct use in obligatory contexts, as there may be few, if any, obligatory contexts for the grammatical construction of interest in the sample. This pitfall can be avoided by carefully controlling the materials used to elicit the language sample and providing contexts specifically designed to elicit a particular grammatical structure (Rice & Wexler, 2001; Rice & Wexler, 1996). However, this may decrease the naturalness and ecological validity of the language sample, which as previously described, is an advantage of language sample analysis.

Another potential drawback of language sample analysis is that it is a relatively time-intensive process. A recent study (Pavelko, Owens, Ireland, & Hahs-Vaughn, 2016) investigated the use of language sample analysis by school-based speech-language pathologists (SLPs) in the United States through a nationwide survey. The researchers found that SLPs cited limited time as the most frequent barrier to utilizing language sample analysis clinically. Although the results of this study showed that SLPs perceive language sample analysis to be a time-consuming process,

some data suggest that shorter samples may be sufficient to yield reliable results (Heilmann, Nockerts, & Miller (2010).

Standardized language assessment. At the heart of language assessment is the question of whether or not a child's language abilities differ significantly from those of their same-age peers. In order to objectively and efficiently answer this question, standardized, norm-referenced tests are used frequently in the diagnosis of language impairment and evaluation of language abilities. Indeed, the utilization of norm-referenced assessments in the diagnosis of language impairment is so prominent that their use is a core component of what is considered the “gold standard” in the assessment of language (Tomblin, Records, & Zhang, 1996).

Advantages of standardized, norm-referenced assessment. First and foremost, standardized, norm-referenced assessments are useful because they provide researchers and clinicians an efficient means for comparing the performance of one child or a group of children to the performance of same-age peers. This is useful for determining if a child or group of children are meeting age-level expectations. Another benefit of standardized, norm-referenced assessments is that they are able to assess a broad assortment of language skills quickly and efficiently. An additional advantage of using standardized assessments is that, because the assessment situation is strictly controlled and the guidelines for assessment explicitly stated, standardized assessments are considered relatively objective measures of language ability (Carrow-Woolfork, 2011).

Disadvantages of standardized, norm-referenced assessment. Despite their widespread use both in research and clinical applications, standardized assessment does have several important limitations. The first and most frequently cited limitation of norm-referenced assessment is that it lacks ecological validity (Costanza-Smith, 2010; Ebert & Scott, 2014). That

is, a child's performance on a standardized assessment of language reflects their ability to perform language-based tasks in an artificial testing situation and does not necessarily directly reflect the child's true language abilities in real-life situations. For example, a child may be unable to name a small, black and white picture of a banana when assessed, but can name a banana while eating her snack at home.

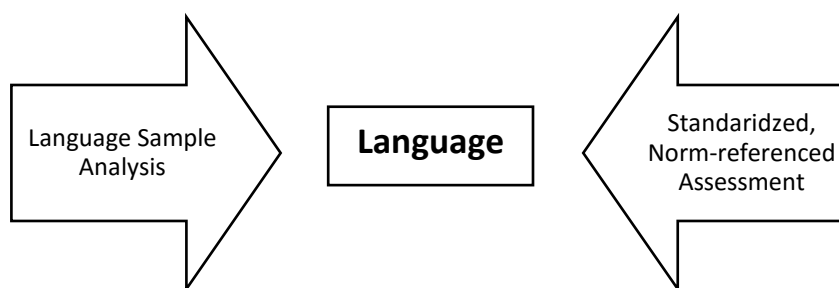
A second limitation of standardized, norm-referenced assessment is that performance and behavioral factors may inhibit a child from demonstrating his or her true language abilities. The majority of standardized assessments require specific behaviors from the child that are not directly related to language functioning. These behaviors often include being seated at a table, attending to the pictorial stimuli for periods of time, responding to the examiner at prescribed times, and guessing when necessary. Children may be unfamiliar or not comply with these behavioral and performance expectations. Due to these expectations placed upon the child, standardized assessment may be poorly suited for some children.

Additionally, the use of standardized assessment with children from linguistically and culturally diverse backgrounds has frequently been called into question (Battle, 2002; De Lamo White & Jin, 2011; Laing & Kamhi, 2003). Because of content and linguistic biases, the scores achieved on standardized, norm-referenced assessments of language may not be valid (Laing & Kamhi, 2003). Content bias refers to bias that is introduced into the assessment when the content of items on the assessment rely on concepts and vocabulary experienced primarily in mainstream culture. Linguistic bias refers to bias that is introduced into an assessment when there is a mismatch between the dialect or language spoken by the child and the dialect or language spoken by the examiner or when a particular language or dialect is required in order to produce an accurate response (Laing & Kamhi, 2003).

Finally, standardized assessments may not be the most effective method for the diagnosis of language deficits in children, despite their widespread use. A study by Plante and Vance (1994) investigated the diagnostic accuracy of four standardized assessments of preschool language with reportedly strong psychometrics. They found that only one of the standardized assessments administered, the SPELT-II (Werner & Kresheck, 1983), reached acceptable levels of accuracy in discriminating between children with and without language impairments. The authors argued that even standardized assessments meeting psychometric criteria might not be adequate for the diagnosis of language impairments when used in isolation, without additional documentation.

Association between language sample analysis and standardized, norm-referenced assessment of language. The relationship between measures taken from language samples and scores on standardized assessments has been studied in both typically developing children and children with language impairments (Condouris, Meyer, & Tager-Flusberg, 2003; Ebert & Scott, 2014; Ukrainetz & Bloomquist, 2002). Because norm-referenced, standardized assessment and the assessment of language through language sample analysis both purport to measure the same construct, language ability, scores obtained from these two different types of assessments should be strongly correlated (see figure 1). However, the results of studies evaluating the relationship between scores obtained on standardized assessments and scores from language sample analysis reveal a more complicated picture.

Figure 1.



Condouris, Meyer, and Tager-Flusberg (2003) investigated the relationship between performance on standardized assessments of language and measures taken from spontaneous speech samples. The children in their study ranged in age from 4 to 14 and all had diagnoses of autism. All of the children in the study were administered the Clinical Evaluation of Language Fundamentals, a standardized assessment of language, and a language sample, gathered while playing with age-appropriate toys. From the language sample, a measure of semantics (i.e., number of different words) and measures of syntax (i.e., MLU and IPSyn) were calculated. After controlling for the effects of non-verbal IQ and age, the authors found that the semantic measures were correlated (i.e., PPVT was correlated with NDW). One of the grammatical measures, MLU, was also correlated with the grammatical subtests of the CELF. Somewhat surprisingly, IPSyn was not found to be correlated with the grammatical subtests of the CELF. The authors of this study hypothesized that the lack of correlation between IPSyn and scores on the CELF were due to autism-specific factors (i.e. the children used a narrow range of grammatical constructions) and that IPSyn may underestimate the grammatical abilities of children with autism.

Another study investigating the relationship between scores obtained on standardized assessment and measures obtained via language sample analysis was conducted by Ebert and Scott (2013). The participants in this study were children referred for a language assessment ranging in age from 6;0 to 12;8, and were separated into an older group and younger group by age. The children were administered an assortment of standardized assessments including the

CELF and language samples were gathered while the child narrated wordless picture books. The researchers found that the correlations between measures taken from the narrative language sample and standardized assessments were much stronger for the children in the younger age range than the older age range. However, even in the younger group, correlations between the standardized assessment scores and language sample measures ranged from essentially no correlation to moderate, statistically significant correlations. Factors that might have influenced these correlations were not explored in this study. However, the researchers hypothesized that lack of correlations seen between many of the language sample variables and standardized assessments could be attributed to task-related differences. Of particular interest was that none of the word-level language sample measures (e.g., NDW) were related to any of the word-level standardized assessments (e.g., PPVT). Grammatical errors in the narrative sample were closely related to scores on the CELF. In this study, the investigators also determined how often the two assessment methods were in agreement with each other in determining the presence or absence of a language impairment. The researchers found moderate overlap in agreement, depending on the standard deviation cut points used, and argued that these methods should be used in conjunction with one another.

A third study of the association between standardized assessment and language sample measures was conducted by Ukrainetz and Bloomquist (2002). This study specifically evaluated four standardized vocabulary tests and three measures taken from a language sample (number of different words, total number of words, and MLU). The children in this study were preschoolers with typical language development. The researchers found that the vocabulary test scores were more strongly correlated with semantic measures (NDW) than non-semantic measures, which they took to be evidence for criterion validity. However, given that from a theoretical standpoint,

measures of semantic skill taken from a language sample should measure the same construct as an expressive vocabulary test (e.g., lexical diversity), some of the correlations were surprisingly low. For example, the correlation between the Expressive Vocabulary Test and NDW in the speech sample was only .48 and was not statistically significant.

Taken together, the findings from these studies suggest that measures taken from language samples are generally correlated with measures obtained from norm-referenced, standardized assessments of language (see Table 1 for a summary of the studies reported here). However, careful examination of the results of these studies reveals that some correlations between standardized assessment scores and language sample analysis measures are not as strong as might be expected given that both assessment methods purport to measure the same constructs. Despite this observation, none of the studies reported here attempted to determine what other factors might account for the variance in the correlations.

Table 1.

Studies on the relationship between performance on standardized assessments of language and measures taken from spontaneous speech samples.

Reference	Sample	Standardized Assessment Measures	Language Sample Analysis Measures	Results
Ebert & Scott (2013)	N=73 6;0-12;8 yrs Referred for language assessment	<ul style="list-style-type: none"> • CELF • PPVT 	<ul style="list-style-type: none"> • MLU • SI • NDW • Omitted bound morphemes • Omitted words • Word-level errors • Utterance-level errors 	<p>Correlations between measures from language sample and standardized assessment were stronger for the younger group than the older group.</p> <p>Moderate overlap of agreement in the identification of students with language impairments.</p>
Condouris, Meyer, & Tager-Flusberg (2003)	N=44 4-14 yrs Diagnoses of autism	<ul style="list-style-type: none"> • CELF • PPVT-3 	<ul style="list-style-type: none"> • MLU • IPSyn • NDW 	<p>Lexical-semantic measures were found to be correlated (i.e., PPVT was correlated with NDW). Grammatical measures were also found to be correlated (i.e., MLU was correlated with grammatical subtests of CELF).</p>

Ukrainetz & Bloomquist (2002)				IPSyn was not correlated with the grammatical subtests of the CELF.
	N=28 3;0-6;0 yrs Typically developing	<ul style="list-style-type: none"> • PPVT-3 • EVT • ROWPVT • EOWPVT-R 	<ul style="list-style-type: none"> • NDW • Total number of words • MLU 	The vocabulary tests were more strongly correlated with semantic measures (NDW) than non-semantic measures (MLU).

Key: VLBW= CELF = Clinical Evaluation of Language Fundamentals-3, PPVT = Peabody Picture Vocabulary Test, MLU=mean length of utterance, SI = Subordination Index, NDW = number of different words, IPSyn = Index of Productive Syntax, EVT = Expressive Vocabulary Test, ROWPVT = Receptive One Word Picture Vocabulary Test, EOWPVT-R = Expressive One Word Picture Vocabulary Test, Revised

Language Sample Analysis and Standardized Assessment in Children Born Preterm

To this date, no studies have been published that explicitly compare performance between standardized, norm-referenced assessments and language sample analysis in preschoolers who were born preterm. Given that children born preterm frequently exhibit attention-related difficulties (Weijer-Bergsma, Wijnroks, & Jongmans, 2008), this population may present a particularly interesting perspective on the relationship between performance on standardized assessments and measures obtained from language sample analysis and how nonlinguistic factors such as attention, hyperactivity, and non-verbal intelligence may influence this relationship. Currently, few studies exist that investigate the conversational language skills of children born preterm in conjunction with results from standardized assessments. A review of the literature resulted in the identification of only two studies that evaluated the language skills of children born preterm through these two different types of assessments and these studies were conducted with children who were older than preschoolers.

Mahurin Smith, Segebart DeThorne, Logan, Channell, and Petrill (2014) evaluated the language skills of school-age children born preterm through both language sample analysis and standardized norm-referenced assessment. Their sample consisted of 57 children born preterm (at or before 32 weeks GA) without significant neurological impairment and a group of full term

peers matched for age, race, gender, and parental education level. The children were seen three times between the ages of 7 and 10, at which time the children were given the CELF-4 and language samples were collected. A correlational analysis of the scores obtained from the language sample and the standardized assessment was not conducted in this study. However, an interesting discrepancy emerged when the authors compared performance between preterm and full term groups on the two different types of language assessments. The children in the preterm group scored significantly below the children in the full term group on the standardized assessment of language, but not on the measures taken from language sample analysis. The authors of this study did not conduct any additional analyses to determine why this discrepancy existed; however, they hypothesized that attentional factors, such as the presence of ADHD, may have influenced the significant preterm-full term difference seen on the standardized assessment scores.

Crosbie, Holm, Wandschneider, and Hemsley (2011) also assessed language via both standardized assessment and language sample analysis and obtained similar results. The participants in this study were 15 school-aged children (mean age 10;6) born at less than 33 weeks GA and a control group of age-matched, full term peers. The children were administered a broad array of neurodevelopmental assessments including the CELF-4. Additionally, a narrative language sample was obtained while the child told a story from a picture book. In this study, the children in the preterm group performed similarly to their full term peers on all of the semantic or syntactic measures obtained from the language samples, which were taken from a narrative sample. The only measure obtained from language sample analysis found to be significantly different between the full term group and preterm group was the number of mazes in the sample, which may indicate higher-level language formulation difficulties for children born preterm. On

omnibus scores from the standardized, norm-referenced assessment (the CELF-4), the children in the preterm group scored, on average, just over one standard deviation below the full term group. When evaluating performance on individual subtests, the children in the preterm group performed more poorly on the formulating sentences and recalling sentences subtests, both of which assess expressive syntax and morphology. These differences were statistically significant.

These two studies provide evidence that children born preterm may perform more poorly when language skill is measured through standardized, norm referenced assessment rather than through language sample analysis. These studies have been conducted with school-age children and the extent to which these children may or may not have received and benefited from intervention was not reported. Studies with younger children born preterm that include standardized assessment and language sample analysis have not been conducted. Further, no studies currently exist for any age of children born preterm that investigate the impact of non-linguistic factors, such as hyperactivity, attention, and non-verbal intelligence, have on standardized assessment results. More research needs to be conducted in this area before conclusions can be drawn on the conversational language skills of children who were born preterm.

Rationale for Current Study

Although standardized assessment of language is an important component of the evaluation of a child's linguistic system, assessing the language skills of a child via language sample analysis provides an ecologically valid assessment of a child's language skills. That is, it offers insight into the functioning of the child's language system in real life situations. Additionally, language sample analysis may be a more sensitive diagnostic indicator of language impairment than standardized assessment in children (Dunn, Flax, Sliwinski, & Aram, 1996).

More needs to be known about the conversational language skills of children who were born preterm in order to broaden the overall picture of the language functioning in this population. Currently, the literature in the area of language outcomes of children born preterm focuses almost exclusively on results of standardized assessments, and those few studies reporting results from language sample analysis have reported only on school-aged children (Crosbie et al., 2011; Mahurin Smith et al., 2014). This study addresses a gap in the literature by evaluating the conversational semantic and grammatical skills of children who were born preterm via language sample analysis.

Secondly, this study seeks to better understand the relationship between performance on standardized assessment and performance when language skill is measured through language sample analysis. Currently, two studies exist that indicate children born preterm perform better when their language skills are evaluated through language sample analysis than when they are evaluated through standardized assessment. This apparent discrepancy may exist for a variety of reasons. One plausible explanation is that performance-related factors such as attention may negatively impact a child's score on standardized assessment. Given that rates of ADHD are very high in children born preterm and deficits in attention and executive functioning are commonly cited (Bhutta et al., 2012; Butcher et al., 2002), this is one plausible explanation.

Another possible explanation for why this discrepancy exists is that language sample analysis assesses only a child's expressive abilities, and standardized assessment usually evaluates both receptive and expressive language. There have been some studies conducted that support the notion that children born preterm have lower receptive language abilities than expressive language abilities (Lewis, Singer, Fulton, Short, Klein, & Baley, 2002; Singer, Siegel, Lewis, Hawkins, Yamashita, & Baley, 2001). However, the two studies previously reported that

measured language skills via both language sample analysis and standardized assessment did not find poorer performance on expressive than receptive measures. More research needs to be conducted to more fully understand these complex relationships.

Research Questions

The primary objective of this study was to investigate the conversational semantic and grammatical skills of preschool-aged children who were born preterm through language sample analysis and standardized, norm-referenced assessment. A secondary objective was to explore the relationship that non-linguistic factors including attention, hyperactivity, and non-verbal intelligence play in standardized assessment outcomes. The following specific research questions were addressed.

1. How do the language skills of preschool-aged children born preterm compare to full term, typically developing children as measured by standardized, norm-referenced assessment, and as measured by language sample analysis?

Hypothesis 1a: The children in the preterm group were predicted to perform more poorly than the children in the full term group when language skill was measured by standardized, norm-referenced assessment. Based on the results of a recent meta-analysis, the Cohen's d effect sizes for the various measures were predicted to be between $-.30$ and $-.82$ (van Noort-van der Spek, Franken, & Weisglas-Kupers, 2012).

Hypothesis 1b: It was predicted that children born preterm would not perform statistically differently from their full term peers on measures of semantic and grammatical language skills obtained from language samples. Although the language skills of preschool-aged children born preterm have not been analyzed via language sample analysis, this prediction was based on previous studies of the conversational and narrative skills of

older, school-aged children born preterm (Crosbie, Holm, Wandschneider, & Hemsley, 2010; Mahurin Smith et al., 2014). These studies found that children born preterm did not perform more poorly than their full term peers on language measures obtained from language sample analysis, such as MLU and NDW.

2. How do the attention skills and non-verbal intelligence of preschool-aged children born preterm compare to full term, typically developing children?

Hypothesis 2: It was predicted that children born preterm would perform more poorly on both parent report of attention problems and hyperactivity and a standardized assessment of attention. This prediction was based on previous studies of attention abilities of children born preterm (Bhutta et al., 2002; Butcher et al., 2002). It was also predicted that the children born preterm would have lower non-verbal intelligence scores (Kerr-Wilson et al., 2012)

3. Are group differences on standardized assessments of language exhibited after controlling for the effects of the non-linguistic factors of attention, hyperactivity, and non-verbal intelligence?

Hypothesis 3: It was predicted that children born preterm would not perform statistically differently from their full term peers on measures of semantic and grammatical language skills obtained from standardized assessments, after controlling for the effects of non-linguistic factors.

Methods

Participants

Children in the preterm group were recruited from an existing database of 92 children previously seen for an assessment of neurodevelopmental outcomes at 30 months of age (Loeb, Imgrund, Lee, & Barlow, in preparation). The children in the preterm group also were participants in a previous study of the effects of patterned orocutaneous stimulation on the feeding skills of preterm infants (Barlow et al., 2014). Inclusionary criteria for this study were birth gestational age before 32 weeks as determined by obstetric ultrasound and clinical examination, tube feedings while in the NICU, head circumference within the 10-90th percentile of mean for post-menstrual age (PMA), neurological examination showing no anomalies for PMA, response to light, sound, spontaneous movements of all extremities, and stable vital signs. Infants with intracranial hemorrhage grades III and IV, neonatal seizures, periventricular leukomalacia, necrotizing enterocolitis, meningitis, sepsis, chromosomal anomalies or craniofacial malformation were excluded from the study.

Thirty children born preterm were recruited to participate in the study. One child recruited for the PT group was unable to complete the assessments due to a significant disability caused by both her preterm birth as well exposure to neurotoxic chemotherapy treatments early in life. The data gathered from this participant were not included in this research because the child was unable to complete the majority of the assessments. Removal of this child from the data set resulted in a total of 29 children in the PT group (18 males, 11 females).

All children in the PT group were between 4;0 and 4;11 years of age at the time of participation. The children were born between 23 and 34 completed weeks of gestation, with a mean gestational age of 31 weeks (SD=19.2). The birthweights of the children in this group

ranged from 610 grams to 2,340 grams, with a mean birthweight of 1,507 grams (SD=461 g).

Socioeconomic status was measured in this study through maternal education level. The mothers of the children who participated in this study came from a variety of education levels (a summary of participant birth, health, and demographic information can be found in table 2). English was the primary language spoken in the home for all participants. One child had minimal exposure to Spanish at home (i.e., exposed to Spanish less than 20% of the time).

Table 2.

Participant characteristics.

	Preterm Mean (SD)	Full Term Mean (SD)	p
Age (years; months)	4;5 (3.1 mo.)	4;4 (3.5 mo.)	.295
Gestational Age (weeks)	31 (18 days)	N/A	
Birthweight (g)	1,507.86 (461)	N/A	
	Preterm n (%)	Full Term n (%)	
Gender			.601
Male	18 (62.1%)	16 (55%)	
Female	11 (37.9%)	13 (44%)	
Maternal Education Level			1.000
Some High School or Less	1 (3.4%)	0 (0%)	
High School/GED	0 (0%)	1 (3.4%)	
Some College/Associate's Degree	9 (31%)	9 (31%)	
Bachelor's Degree	11 (37.9%)	12 (41.4%)	
Graduate Degree	8 (27.6%)	7 (24.1%)	
Race			.569
White	26 (89.6%)	24 (82.8%)	
Black/African American	1 (3.4%)	2 (6.9%)	
Asian	0 (0%)	0 (0%)	
American Indian/Alaskan Native	0 (0%)	0 (0%)	
More than one race	2 (6.9%)	3 (10.3%)	
Ethnicity			.455
Hispanic/Latino	3 (10.3%)	5 (17.2%)	
Not Hispanic/Latino	26 (89.6%)	24 (82.8%)	

The children in the FT group were recruited from local preschools in the Kansas City metropolitan area and through fliers placed in community locations such as libraries, doctor's offices, and community posting boards. Because factors such as gender and socioeconomic status can influence language development, efforts were made to ensure that the FT and PT groups were demographically similar to one another. The PT and FT groups contained similar gender and racial/ethnicity ratios. Additionally, the two groups had similar levels of maternal education, which was used as a proxy measure of socioeconomic status (see Table 2). Independent samples T-tests indicated that there were no statically significant group differences on the reported demographic variables (see Table 2).

Eligibility criteria for the FT group included birth after 37 completed weeks of gestation, no history of hearing loss, language or other developmental delay or disorder, behavioral or attention disorder, or a vision impairment that was not correctible by the use of glasses. This information was obtained via a parent questionnaire (see Appendix A). Eligibility testing also was conducted to determine which potential participants qualified for the FT group. Only children with nonverbal IQ scores and composite language scores above 85 were considered eligible for the FT comparison group. A total of 32 children were recruited for the FT group. Three children did not meet eligibility criteria, resulting in a comparison group of 29 preschoolers born full term.

Procedures

In order to address the study aims, the children attended a single assessment session, which took place at a quiet location in the community such as a study room in a local library or a clinic room inside the Schiefelbusch Speech-Language-Hearing clinic at the University of Kansas. The primary investigator ensured that the room was free from distractions. Parents were

allowed to attend the session with their child but were asked not to interact with their child during the assessments. The assessment session lasted approximately 90 minutes, which included short breaks between assessments as well as a longer snack break. The children were given a small token (i.e., a sticker to be placed on a sticker chart) after the completion of each task in order to increase child compliance with the assessment tasks. The assessments were administered in a randomized order to ensure order effects would not influence the assessment results.

Evaluation of language: Language sample analysis. Children's conversational language skills were assessed through analysis of language samples. The language samples were obtained during free play with the researcher and utilized a standard set of toys, which were presented to the children in a random order. Three sets of toys were used: 1) a toy barn set with accompanying animals, people, and props associated with a barn, 2) dolls with accompanying accessories, clothes, and diapers, and 3) fairytale-themed puppets including a dragon, prince, princess, and knight (Appendix B). These items were selected because of their developmental appropriateness, high-interest level for preschool-aged children, and their ability to provide contexts for the elicitation of a wide range of language structures. The examiner implemented Leadholm and Miller's (1994) recommendations for collecting a language sample by following the child's lead and using primarily open-ended prompts and questions. The language samples were video and audio recorded. The language samples were transcribed by the primary investigator using conventional Systematic Analysis of Language Transcripts (SALT) procedures (Miller & Chapman, 2007). So that each analysis was of an equal length, only the first 100 complete and intelligible utterances were used in the analyses, which is a commonly used standard in the analysis of language samples (Cole, Mills, & Dale, 1989; Heilmann, 2010; Miller, 1981).

In order to assess conversational language skills, measures of semantic ability and grammatical ability were obtained from the language samples. Semantic measures included the number of different word roots used by the child (NDW) and a semantic analysis. NDW has been used to differentiate children with language impairment from children with typical language skills and is considered a valid measure of lexical diversity (Malvern & Richards, 2012; Watkins, Kelly, Harebers, & Hollis, 1995). NDW also has been found to correlate with standardized measures of expressive vocabulary (Ukrainetz & Bloomquist, 2002). A semantic analysis also was conducted to measure semantic skill. The use of lower-frequency words such as “success” and “collide” can add precision and depth to the conversations of young children, and has been considered a marker of increased semantic ability (Beck, McKeown, & Kucan, 2013). In order to conduct the word-frequency analysis, each of the different word roots spoken by the individual child were tallied. From this list of word roots, the high-frequency words commonly used by young children were removed, leaving only lower-frequency words. High-frequency words were identified by using the MacArthur-Bates database of words produced commonly by toddlers (Fenson et al., 1991). Additionally, proper nouns, sound effects, and numbers were removed from the list. The remaining lower-frequency words were tallied for each child, resulting in a semantic analysis score. This procedure was adapted from Mahurin Smith, DeThorne, Logan, Channell, and Petrill (2014).

Three distinct measures of syntactic and morphological skills were also obtained from the language samples 1) mean length of utterance, 2) percentage correct use of finite verb morphology in obligatory contexts, and 3) Developmental Sentence Score (DSS). These measures were selected because they each offer unique insight into different aspects of a child’s developing grammatical skills.

Mean length of utterance (MLU) in morphemes was calculated for each of the children by calculating the total number of morphemes in the language sample and dividing that number by the total number of utterances used by the child (Brown, 1973). Although there are limitations to using MLU as the only measure of syntax (Eisenberg, Fersko, & Lundgren, 2001), MLU is a commonly used gross measure of syntactic ability and is a valuable measure for assessing children's productive language skills, especially when used in conjunction with other measures (Leonard & Finneran, 2003). MLUs were computed using SALT software.

The second measure was based on the children's production of finite verb morphology in obligatory contexts (Goffman & Leonard, 2000). Difficulty with finite verb morphology is a key indicator of language impairment, and the assessment of finite verb morphology is considered a particularly robust method for the identification of LI in young children (Rice & Wexler, 2001). Goffman and Leonard's (2000) methods for calculating percentage correct use of finite verb morphology in obligatory contexts was utilized in this study. The finite verbs and morphemes under investigation included past tense –ed, third person singular –s, and the copula and auxiliary verbs *is*, *are*, *am*, *was* and *were*. For each child, percent correct use of finite verb morphology in obligatory contexts was determined by dividing the number of correct uses by the total number of obligatory contexts and multiplied by 100 to yield percent correct. In order to ensure that an adequate number of obligatory contexts were elicited, the examiner periodically used particular prompts to elicit third person present –s and past tense –ed. For example, the researcher used the prompt “tell me what your ____ (e.g., baby, car, boy) does?” (adapted from Rice & Wexler, 2001) to elicit third person singular –s. To attempt to elicit regular past tense –ed, the examiner periodically focused the play on actions that are conventionally named in English with a regular verb (e.g., walk, hug, kiss, play, jump) and prompted with “what happened?” to elicit past tense

–ed (adapted from Rice & Wexler, 1996). Both of these prompts were utilized by the researcher a minimum of five times each throughout collection in order to provide adequate obligatory contexts for the production of finite verb morphology. In order to determine that the participants were able to produce the phonemes required for the marking of tense, the participants were administered the Phonological Probe from the Test of Early Grammatical Impairment (TEGI) (Rice & Wexler, 2001). All of the participants in both the PT and FT groups passed the probe.

3) The third measure of syntax was the child's Developmental Sentence Score (DSS) (Lee, 1974). DSS is a procedure for estimating a child's syntactic abilities and includes information on eight syntactic categories: indefinite pronouns, personal pronouns, main verbs, secondary verbs, negatives, conjunctions, interrogative reversals, and -wh questions. For every utterance, the child is given a point value between 1 and 8 for each of the syntactic categories used in the utterance. Lower point values are assigned to early-acquired syntactic constructions and higher point values assigned for later-acquired syntactic constructions. For example, uninflected main verbs (e.g., see, kick, go) are awarded one point and more complex constructions such as have + been + verb + ing (e.g., have been walking) are awarded eight points. Additionally, each utterance is awarded one point if the utterance is completely correct by adult standards of grammar. Only complete, (i.e., the utterance contains a subject and verb), unique, and intelligible utterances are scored using this procedure. In order to be eligible for a DSS analysis, a language sample must contain at least 50% complete, unique and intelligible utterances. The DSS was calculated for each of the child's utterances in the sample, and then averaged, yielding the overall DSS. Although other methods for quantifying the grammatical skills of children exist (e.g., IPSyn), DSS was selected for use in this study because it provides extensive information about a child's grammatical skills across many different grammatical

categories and has been shown to be effective in differentiating between children with and without language impairments (Paul, 1996). Despite being over 40 years old, DSS maintains its utility as a measure of syntactic ability in both clinical and research contexts (Hughes, Fey, & Long, 1992). The procedure continues to be used widely for both children who are typically developing and children with language impairments, and has even been adapted for non-English languages (Miyata et al., 2013; Mahurin Smith et al., 2014; Souto, Leonard, & Deevy, 2014).

Transcription and coding reliability analyses. In order to ensure that the data obtained from the language samples were reliability transcribed and coded and that the language sample analysis variables were reliably calculated, a series of reliability analyses were conducted. The following formula was used to determine transcription reliability and inter-rater reliability for DSS, percent correct use of finite verb morphology in obligatory contexts, and the semantic analysis score:

$$\frac{\text{\# of agreements}}{\text{\# of agreements} + \text{\# of disagreements}} \times 100$$

Twenty percent of the language samples were transcribed independently by a second researcher. After the independent researcher concluded her transcription, the two transcriptions were compared on a morpheme-by-morpheme basis. The number of agreements and disagreements were calculated. For example, if the primary researcher transcribed “That/’s my baby” and the independent researcher transcribed “That/’s your baby”, this would yield three agreements and one disagreement. Inter-rater reliability ranged from 92.3 - 97.4% (mean = 94.7%) in the PT group and from 94 - 96.3% (mean = 95.4) in the FT group.

Inter-rater reliability was also calculated for the semantic analyses on 20% of the participants. An independent researcher reviewed each of the word roots produced by the child

and recorded her judgment on whether each word should be retained or rejected from the overall tally. Point-to-point reliability ranged from 98.4 - 100% (mean = 99.4%) in the PT group and 97.6 - 99.3% (mean = 98.7%) in the FT group

Likewise, the coding of correct use in obligatory contexts was calculated independently by a second researcher for 20% of the transcripts on an utterance-by-utterance basis. For each utterance, instances of obligatory contexts for third person singular –s, past tense –ed, copulas and auxiliary verbs were recorded by an independent researcher. The independent researcher also recorded if the child used the obligated finite verb morphology correctly. The independent researcher's responses were compared to the original sample on a point-by-point basis and the number of agreements and disagreements were then calculated. Inter-rater reliability ranged from 87.5 - 97.7% (mean = 94.5%) in the PT group and 89.4 - 93.5% (mean = 91.7%) in the FT group.

For DSS, the primary researcher completed five training DSSs under the guidance of a DSS expert in order to become proficient with scoring. A DSS score was also calculated for an additional 20% of the transcripts by the DSS expert in order to determine inter-rater reliability for the DSS scores. The independently scored DSSs were compared to the original scores on a category by category basis. The number of agreements and disagreements were then calculated. For example, if the sentence “she is sleeping” was scored as a 2 for the personal pronoun category, a 1 for the main verb category, and a 1 for being correct by adult standards, but the second rater scored the sentence 1, 1, 1 respectively, this would yield one disagreement and two agreements. Inter-rater reliability ranged from 93-98.5% (mean=94.6%) in the PT group and 92-96.2% (mean=93.9%) in the FT group.

Evaluation of Language: Formal Standardized Assessment. The Clinical Evaluation of Language Fundamentals-Preschool- 2 (CELF-P2) (Wiig, Secord, & Semel, 2004) is a formal standardized language assessment used to evaluate a broad range of language skills of children in the preschool age range and is widely used in the clinical diagnosis of language impairment. The administration of this assessment results in a receptive language standard score, an expressive language standard score, and a core language standard score. Two additional composite scores can be obtained: 1) language content index (i.e., a measure of semantic skill) and 2) language structure index (i.e., a measure of morphosyntactic skill).

The CELF-P2 is a psychometrically strong assessment tool with high levels of both reliability and validity (Friberg, 2010; Semel, Wiig, & Secord, 2004). Reliability refers to an assessment's accuracy, consistency, and stability and is often assessed by determining test-retest stability, internal consistency, and inter-rater consistency (McCauley & Swisher, 1984). The validity of a test refers to the extent that an assessment tool measures what it was intended to measure, as supported by theory and research. Validity is often assessed through the conduction of item-analysis and by determining concurrent validity (McCauley & Swisher, 1984). Friberg (2010) used 11 criteria to assess the psychometric properties, including validity and reliability, of several standardized assessments available for young children, including the CELF-P2. The CELF-P2 met 10 of the 11 criteria: test purpose defined, tester qualifications indicated, procedures explained, adequate sample size, sample clearly defined, evidence of item analysis, measures of central tendency, concurrent validity, test/retest reliability, and inter-examiner reliability. Based on this information, scores obtained on the CELF-P2 can be considered both valid and reliable.

Although the CELF-P2 was given in its entirety to the participants for eligibility purposes, the expressive language subtests designed to assess expressive semantics and grammar were of particular importance in order to address the aims of this study. These subtests were: 1) Word Structure, 2) Expressive Vocabulary, and 3) Recalling Sentences. The Word Structure subtest was designed to evaluate a child's use of grammatical rules in a sentence-completion task. In this test, the child is asked to look at a picture and complete a sentence using a targeted word structure (e.g., irregular past tense, present progressive verb, reflexive pronoun). The Expressive Vocabulary subtest was designed to measure expressive semantics and requires the child to label pictures of people, objects, and actions. Finally, the Recalling Sentences subtest was designed to evaluate a child's ability to repeat sentences of increasing length and grammatical complexity verbatim. Although sentence repetition tasks are commonly used as a measure of a child's grammatical abilities on standardized assessments such as the CELF-P2, these tasks have been noted to also assess other aspects of language and cognition (Archibald & Joanisee, 2009).

Attention. In order to assess attention, two measures were used: 1) a parent report measure of ADHD symptoms and 2) results from a standardized performance-based attention assessment. The use of multiple measures, including the results of both parent report and standardized assessment, is recommended in the evaluation of attention and hyperactivity deficits in young children (American Academy of Pediatrics, 2011; Mahone & Schneider, 2012).

The Behavior Assessment for Children, Second Edition (BASC-3)(Reynolds & Kamphaus, 2015) was used to assess inattention and hyperactivity/impulsivity symptoms. The BASC-3 is a broad-based assessment designed to assess a wide range of behavior problems in children, including attention and hyperactivity problems. The BASC-3 contains several scales,

including the parent rating scale (PRS), which allows for the parent report of behaviors in their children. Evaluation of the psychometric properties of the BASC-3 indicates strong reliability and validity (Reynolds & Kamphaus, 2015) as evidenced by good to excellent internal reliability coefficients and test-retest correlations. The rating scale was completed by the child's primary caregiver and took approximately 15 minutes to complete. Although the BASC-3 was administered in its entirety, only the scales assessing inattention and hyperactivity were used in the analyses for this study.

The Conners Kiddie Continuous Performance Test 2nd Edition (Conners K-CPT 2) (Conners, 2015), a computerized performance-based assessment of attention deficits in children ages 4-7, was administered to each child. The assessment uses pictures of common items that are familiar to young children (e.g., boat, ball, etc.). The children were trained to respond by clicking a computer mouse after seeing a picture of a target (all objects except the ball) and to refrain from responding after seeing a non-target (a ball). The Conners K-CPT 2 measures performance in four areas related to attention deficits: 1) inattentiveness, 2) impulsivity, 3) sustained attention, and 4) vigilance. Although the administration of the Conners K-CPT 2 results in the generation of many variables, only the T-score for omissions was used for the purposes of the present study. The omission T-score is derived from the number of targets that the child did not respond to and is a measure of inattentiveness (Conners & Staff, 2001). This measure was determined to be the most suitable variable for use in this study because it could be used as a standalone measure of inattentiveness. The other variables obtained from this test need to be interpreted in tandem with other variables in order to be of diagnostic value. The entire assessment was administered in 7.5 minutes. In order to ensure that all children were familiar with the targets used on the Conners K-CPT 2, a short training period was utilized at the

beginning of each assessment. All children were able to adequately demonstrate the ability to click the mouse in response to targets and refrain from clicking the mouse in response to non-targets.

Other assessment measures. In order to assess non-verbal intelligence, the Kaufman Brief Intelligence Test, 2nd Edition (KBIT-2) was administered (Kaufman & Kaufman, 2004). The KBIT-2 is a standardized assessment of intelligence. Only the non-verbal portion of this assessment was administered.

A hearing screening was conducted with a calibrated audiometer and over the ear earphones at 20dB at 1,000, 2,000, and 4,000Hz (American Speech-Language-Hearing Association, 1997). All but one child from the PT group passed the hearing screening. This child could not be reliably conditioned to respond to the tones and was not compliant with wearing the earphones. Her parents noted no concerns with hearing and she had recently passed a hearing screening conducted by her pediatrician. In the FT group, all but one child also passed the hearing screening. Her parents noted no concerns with hearing and the child passed a hearing test provided by an audiologist shortly after participating in this study. The parents of both children denied recent occurrence of ear infection or head cold. The data obtained from these children were included in this study because their parents provided documentation of adequate hearing.

Results

Statistical Analyses

The primary objective of this study was to investigate the conversational semantic and grammatical skills of preschool-aged children who were born preterm through language sample analysis and standardized, norm-referenced assessment. A secondary objective was to explore

the relationship that non-linguistic factors including attention and non-verbal intelligence play in standardized assessment outcomes.

In order to address the research questions, the means and standard deviations were calculated for each of the dependent variables of interest (see Table 3). Additionally, the percentage of children scoring below the average range was calculated for each of the language variables (see Table 4). Cutoff scores for each of the variables were obtained via a variety of methods. For the standardized assessment variables obtained from the subtests of CELF-P2, scaled scores below 7 were considered below the average range, which corresponds with one standard deviation below the mean. The average range cutoff for NDW was obtained from the SALT reference database. In the SALT reference database, the average NDW for children ages 4;0 to 4;11 was 149 with a standard deviation of 35.24. Thus, an NDW below 114 corresponds with one standard deviation below the mean; scores below this cutoff were considered below the average range. For MLU in morphemes, the Rice et al. (2010) norms were used to determine average range cutoff scores. MLUs below 3.81 and 3.96 (one standard deviation below the mean in their normative sample) were considered below the average range for children 4;0-4;5 and 4;5-4;11 respectively. Lee's (1974) normative data was used to determine a cutoff score for the DSS (pg. 167). Scores below the 10th percentile were considered below the average range, as per Lee's recommendations. Finally, for percent correct use of finite verb morphology in obligatory contexts, a cutoff score of 85% was selected based on Goffman and Leonard's (2000) normative data.

Table 3.
Descriptive Statistics for Dependent Variables of Interest.

	Preterm Mean (SD)	Range	Full Term Mean (SD)	Range
Language Sample Analysis				
Semantic: NDW	130.34 (19.34)	85 - 184	142.86 (14.93)	117 - 171
Semantic: Semantic Analysis Score	35.90 (9.54)	20 - 57	41.72 (7.21)	30 - 54
Grammatical: MLU	3.897 (.58)	2.41 - 5.09	4.433 (.49)	3.23 - 5.53
Grammatical: DSS	6.824 (.90)	4.90 - 8.98	7.419 (.96)	5.40 - 9.36
Grammatical: % correct use of finite morphology	81.21 (12)	46.88 - 100	93.21 (6.2)	77.14 - 100
Standardized Assessment of Language				
Semantic: CELF P2 Expressive Vocabulary	10.62 (2.15)	4-14	10.90 (2.01)	7-16
Grammatical: CELF P2 Word Structure	10.41 (2.26)	5-15	11.14 (1.73)	8-14
Grammatical: CELF P2 Recalling Sentences	9.17 (2.43)	3-12	10.79 (1.90)	7-16
Attention/Hyperactivity				
BASC-3 Attention Problems Scale	47.92 (6.25)	38 - 72	49.48 (5.81)	36 - 58
BASC-3 Hyperactivity Scale Conners K-CPT2	48.79 (8.21)	37 - 80	52.31 (7.59)	39-69
Non-verbal Intelligence				
KBIT-2	100.24 (13.09)	64 - 124	103.83 (5.75)	88 - 115

Table 4.

Percentage of children scoring below the average range on language assessment variables.

	Preterm n (%)	Full Term n (%)
Language Sample Analysis		
NDW	4/29 (13.8%)	0/29 (0%)
MLU	14/29 (48.3%)	2/29 (6.9%)
DSS	6/25 (24%)	3/29 (10.3%)
% correct use of finite morphology	17/29 (58.6%)	4/29 (13.8%)
Standardized Assessment		
CELF P2 Word Structure	2/29 (6.9%)	0/29 (0%)
CELF P2 Expressive Vocabulary	1/29 (3.4%)	0/29 (0%)
CELF P2 Recalling Sentences	5/29 (17.2%)	0/29 (0%)

Analysis of Variance (ANOVAs)

To determine if differences between the PT and FT groups were present, a series of ANOVAs were conducted to compare performance of the preterm group to the full term group on each of the variables obtained from the language samples, CELF-P2, and the non-linguistic factors. A modified Bonferroni adjustment was used to determine the statistical significance of the p-values for all sets of ANOVAs. For this type of adjustment, the alpha levels were adjusted based on the number of comparisons conducted for each of the outcomes of interest. This adjustment was used in order to preserve power while ensuring the probability of committing a Type I error remained below .05. Effect sizes were calculated for each of the variables of interest to evaluate practical significance. F-values from the ANOVAs were converted directly into Cohen's d values using an effect size calculator. For the purposes of the present study, the following guidelines for interpreting effect sizes were used: small effect $d = .2$, medium effect $d = .5$, large effect $d = .8$ (Cohen, 1977).

Research Question 1

Standardized assessment of language results. A one-way between subjects ANOVA was conducted to compare performance on the expressive language subtests of the CELF-P2, which includes one subtest assessing expressive semantics (i.e., Expressive Vocabulary subtest) and two subtests assessing expressive grammar (i.e., Word Structure and Recalling Sentences subtests). For the semantic measure, there was not a statistically significant difference between the PT and FT groups [$F(1,56) = .256$, $p = 0.617$]. The magnitude of the effect was negligible ($d = .135$). For the grammatical measures, there was not a significant difference between groups on performance on the CELF-P2 Word Structure subtest [$F(1,56) = 1.88$, $p = 0.176$], with a small effect size ($d = .366$). Statistically significant group differences were seen on the CELF-P2

Recalling Sentences subtest [$F(1,56) = 7.99, p = 0.007$]. The magnitude of the effect was medium ($d = .755$), indicating both statistical and practical significance for this measure. These results are summarized in Table 5.

Table 5.
One-way Analysis of Variance: CELF-P2

		Sum of Squares	df	Mean Square	F	Sig.
CELF-Word Structure Scaled	Between Groups	7.603	1	7.603	1.880	.176
	Within Groups	226.483	56	4.044		
	Total	234.086	57			
CELF-Expressive Vocabulary Scaled	Between Groups	1.103	1	1.103	.256	.615
	Within Groups	241.517	56	4.313		
	Total	242.621	57			
CELF-Recalling Sentences Scaled	Between Groups	38.086	1	38.086	7.991	.007
	Within Groups	266.897	56	4.766		
	Total	304.983	57			
ELI Standard	Between Groups	220.155	1	220.155	2.320	.133
	Within Groups	5314.000	56	94.893		
	Total	5534.155	57			

Language sample analysis results. Likewise, a one-way between subjects ANOVA was conducted to compare performance between the FT and PT groups on the measures obtained from language sample analysis. Statistically significant group differences were found for all measures obtained from the language sample analyses. For the semantic measures, significant group differences were found for both NDW [$F(1,56) = 7.612, p = .008$] and the semantic analysis score [$F(1,56) = 6.883, p = .011$]. The magnitudes of the effects were medium ($d = .754$ and $d = .70$, respectively). On the grammatical measures, statistically significant group differences were found for MLU [$F(1,56) = 14.508, p < .001$], DSS [$F(1,52) = 5.452, p = .023$], and percent correct use of finite verb morphology [$F(1,56) = 22.757, p < 0.001$]. The effect sizes for MLU and percentage correct use of finite verb morphology were large ($d = 1.02$ and $d = 1.28$, respectively). The effect size for DSS was medium ($d = .65$). It should be noted that four children

from the PT group did not have DSS scores, because their language samples did not meet criteria for computing a DSS (i.e., less than 50% of the utterances contained a subject and verb). These results are summarized in Table 6.

Table 6.

One-way Analysis of Variance: Language Sample Analysis

		Sum of Squares	df	Mean Square	F	Sig.
Number of Different Words	Between Groups	2271.879	1	2271.879	7.612	.008
	Within Groups	16714.000	56	298.464		
	Total	18985.879	57			
Semantic Analysis Score	Between Groups	492.431	1	492.431	6.883	.011
	Within Groups	4006.483	56	71.544		
	Total	4498.914	57			
Mean Length of Utterance	Between Groups	4.164	1	4.164	14.508	<.001
	Within Groups	16.071	56	.287		
	Total	20.235	57			
Developmental Sentence Score	Between Groups	4.755	1	4.755	5.452	.023
	Within Groups	45.357	52	.872		
	Total	50.113	53			
% correct oblig finite verb morphology	Between Groups	.209	1	.209	22.757	<.001
	Within Groups	.514	56	.009		
	Total	.722	57			

Research Question 2

Attention/hyperactivity and non-verbal intelligence results. To determine if group differences were present for the variables measuring the non-linguistic skills of attention, hyperactivity, and non-verbal intelligence a one-way between subjects ANOVA was conducted. Group differences were not seen for either parent report of hyperactivity [$F(1,56) = 2.868$, $p = .096$] or attention problems [$F(1,56) = .914$, $p = .343$] on the BASC-3 parent rating scales. Effect sizes for these measures were small ($d = .45$ and $d = .26$). On the performance measure of attention, omission T scores obtained from the Conners K-CPT2, group differences were not observed [$F(1,55) = .000$, $p = .984$]. Finally, group differences were not observed for scores on

the non-verbal intelligence subtest of the KBIT [$F(1,56) = 1.825, p = .182$]. These results are summarized in Table 7.

Table 7.

Analysis of Variance: Attention and Hyperactivity

		Sum of Squares	df	Mean Square	F	Sig.
BASC Hyperactivity	Between Groups	179.379	1	179.379	2.868	.096
	Within Groups	3502.966	56	62.553		
	Total	3682.345	57			
BASC Attention Problems	Between Groups	33.379	1	33.379	.914	.343
	Within Groups	2044.207	56	36.504		
	Total	2077.586	57			
Conners K-CPT2	Between Groups	.106	1	.106	.000	.984
	Within Groups	13934.034	55	253.346		
	Total	13934.140	56			
KBIT Non-verbal	Between Groups	186.483	1	186.483	1.825	.182
	Within Groups	5721.448	56	102.169		
	Total	5907.931	57			

Research Question 3

Impact of non-linguistic factors. A group difference for only one standardized assessment measure of language (i.e., Recalling Sentences subtest of the CELF-P2) was identified through ANOVA. In order to determine the role that non-linguistic factors including attention, hyperactivity and non-verbal intelligence play in standardized assessment outcomes, a series of analyses of covariance (ANCOVAs) were utilized. If group differences were reduced or negated after the controlling for these factors, it would support the hypothesis that differences seen between the preterm and full term groups on the Recalling Sentences scores could be largely attributed to non-linguistic factors. The results of the ANCOVA indicated that parent report of hyperactivity did not account for a significant amount of the observed group differences in Recalling Sentences scores. Parent report of attention problems did account for a significant

amount of the group difference in Recalling Sentences scores, but the group difference remained after accounting for attention problems. A performance-based measure of attention (i.e., Conners K-CPT) did not account for a significant amount of the observed group difference in Recalling Sentences scores. Likewise, non-verbal intelligence did not account for a significant amount of the observed group difference in Recalling Sentences scores. These results are summarized in Tables 8, 9, 10, and 11.

Table 8.

Analysis of Covariance: BASC-3 Hyperactivity

Dependent Variable: CELF-Recalling Subtest Scaled

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	47.867 ^a	2	23.934	5.120	.009
Intercept	214.425	1	214.425	45.868	.000
BASC Hyperactivity T	9.781	1	9.781	2.092	.154
Group	45.017	1	45.017	9.630	.003
Error	257.115	55	4.675		
Total	6085.000	58			
Corrected Total	304.983	57			

a. R Squared = .157 (Adjusted R Squared = .126)

Table 9.

Analysis of Covariance: BASC-3 Attention Problems

Dependent Variable: CELF-Recalling Subtest Scaled

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	56.484 ^a	2	28.242	6.251	.004
Intercept	180.987	1	180.987	40.058	.000
BASC Attention Problems T	18.398	1	18.398	4.072	.048
Group	44.426	1	44.426	9.833	.003
Error	248.499	55	4.518		
Total	6085.000	58			
Corrected Total	304.983	57			

a. R Squared = .185 (Adjusted R Squared = .156)

Table 10.

Analysis of Covariance: Conners K-CPT2

Dependent Variable: CELF-Recalling Subtest Scaled

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	35.922 ^a	2	17.961	3.835	.028
Intercept	419.829	1	419.829	89.636	.000
Omissions	3.552	1	3.552	.758	.388
Group	32.428	1	32.428	6.924	.011
Error	252.921	54	4.684		
Total	6049.000	57			
Corrected Total	288.842	56			

a. R Squared = .124 (Adjusted R Squared = .092)

Table 11.

Analysis of Covariance: Non-verbal IQ

Dependent Variable: CELF-Recalling Sentences Scaled

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	52.159 ^a	2	26.080	5.673	.006
Intercept	13.190	1	13.190	2.869	.096
Nonverbal Standard	14.073	1	14.073	3.062	.086
Group	29.233	1	29.233	6.359	.015
Error	252.823	55	4.597		
Total	6085.000	58			
Corrected Total	304.983	57			

a. R Squared = .171 (Adjusted R Squared = .141)

Discussion

Language deficits in children born preterm have been well-documented in the literature; however the majority of studies on language development in this population have focused nearly exclusively on the results of standardized assessment. As a result, little is known about the conversational language abilities of children born preterm as measured by language sample

analysis. In comparison to a group of age and SES-matched peers, the children in the preterm group performed more poorly on all measures of language development obtained from the language samples. These measures included both measures of semantic ability (e.g., NDW and the semantic analysis score) as well as grammatical ability (MLU, DSS, and % correct use of finite verb morphology).

Furthermore, the magnitudes of the effects for each of these variables were moderate to large, indicating that group differences exhibited on the ANOVAs are of practical significance. The measures of grammar tended to have larger effect sizes than the measures of semantic ability, indicating that preschoolers born preterm may have more difficulty with grammatical aspects of language than semantic. The one exception to this pattern was the DSS, which had a medium effect size. However, DSSs were unable to be calculated for four of the children from the PT group because 50% of their utterances did not contain a subject and a verb (and therefore did not meet criteria for calculating a DSS). A subjective analysis of these children's language samples showed that they had significant grammatical difficulties.

The results from the language sample variables indicate consistent difficulty with conversational language skills, especially in the area of grammar. When compared to normative data, nearly half of the children in the preterm group had MLUs below the average range, over half did not meet age-level expectations for production of finite verb morphology, and nearly a quarter had DSS scores that were below the average range. From these data, it can be concluded that expressive language impairments were evident in many of the children born preterm when language was measured through language sample analysis.

In contrast, and not as predicted, only 3 out of the 29 (10.3%) children in the PT group had Expressive Language Index (ELI) scores on the CELF-P2 that were below the average

range. Thus, the majority of children who showed deficits on the language sample measures did not score below the average range on the standardized assessment of language. All three of the children who had ELI scores in the below average range also scored below the average range on at least two of the language sample analysis measures.

The present study identified group differences between the PT and FT groups on one of the standardized assessment subtests of the CELF-P2. The Recalling Sentences subtest of the CELF-P2 was the only standardized assessment of expressive language for which group differences were present. A similar result was present in Foster-Cohen's (2010) study of children born preterm who were preschoolers; however, they found a small effect size; whereas in the current study, a medium effect size was present. Sentence repetition tasks, such as that used in the Recalling Sentences subtest, have been cited as a particularly robust assessment of language impairment in children, with some data suggesting that sentence repetition is the best single test for differentiating between children with and without language impairments (Conti-Ramsden, Botting & Faragher, 2001). Although sentence repetition tasks are used commonly in both clinical and research settings, the linguistic and cognitive abilities being measured by this type of assessment are not yet fully understood. Klem et al. (2015) argues that sentence repetition is a multi-faceted task that taps into "virtually all" aspects of language.

An analysis of the role that non-linguistic factors played in the group differences on the CELF-P2 Recalling Sentences scores indicated that of the non-linguistic factors attention, hyperactivity, and non-verbal intelligence, only parent report of attention problems accounted for a significant portion of the observed group difference. Although adding parent report of attention problems into the model accounted for a significant amount of the group difference, the group difference remained statistically significant. This finding supports the hypothesis that non-

linguistic factors such as attention may contribute to performance on standardized tests such as recalling sentences tasks in children born preterm, but does not wholly account for their poorer performance.

The majority of studies of language development in children born preterm and a meta-analysis of these studies indicate consistent group differences when language development is measured via standardized assessment (Casiro et al., 1990; Foster-Cohen et al., 2007; Magill-Evans et al., 2002; van Noort-van der Spek, et al., 2012). Given that studies of language abilities in children born preterm have consistently found that these children perform more poorly than their full term peers when language skills are measured through the use of standardized assessment, the lack of group differences seen between the PT and FT groups on the majority of standardized assessment measures was somewhat surprising. There are several possible explanations for this finding.

First, children born preterm are a very heterogeneous group. The characteristics of children born preterm can vary greatly from study to study. The children who participated in the current study may have been different from the participants in other studies of language skills in meaningful ways that influenced the results of this study. First, the mean gestational age of the children in the present study was 31 completed week of gestation. Several other studies that have documented language deficits in children born preterm have utilized cohorts of children born preterm with a lower mean gestational age. The large cohort of preschoolers born preterm in the Foster-Cohen et al. (2010) study had a mean gestational age of 27.8 weeks. The mean gestational ages of the Mahurin Smith et al. (2014) study was 29.8 weeks and the mean gestational age was not reported for the Crosbie et al. (2010) study. Neurodevelopmental outcomes tend to decrease linearly as gestational age at birth decreases (Curry, Pfeiffer, Slopen, & McVeigh, 2012), thus, a

difference in mean gestational age may have resulted in a different finding. However, results from a systematic review of neurodevelopmental outcomes suggests that even children born late preterm (i.e., 34-36 weeks of gestation) perform less favorably on neurodevelopmental assessment measures than their full term counterparts (McGowan, Alderice, Holmes, & Johnston, 2011).

Socioeconomic differences between the sample used in the present study and previous studies of language outcomes of children born preterm may have also contributed to the lack of group differences seen on most of the standardized assessment measures. Studies of language development in children born preterm frequently use maternal education level as a gross measure of socioeconomic status. In the current study, all but one of the mothers had completed some amount of college, and over 65% had completed a bachelor's or graduate degree. In the large Foster-Cohen cohort of preschoolers born preterm, 40% of the mothers had not completed secondary school. Given that rates of language impairment have been shown to vary significantly among different socioeconomic levels, this factor may have skewed the current sample toward more favorable language outcomes.

Another possible explanation for the lack of group differences exhibited on most standardized assessment measures is access to and utilization of early intervention services. Early intervention has been shown to have a positive impact on the language and developmental outcomes of children at risk for poor language development, including children who were born preterm (Orton, Spittle, Doyle, Anderson, & Boyd, 2009). The Program for Infants and Toddlers with Disabilities (Part C of the Individuals with Disabilities Education Act) is a federal program that assists individual states in establishing early intervention programs for young children and their families. In the present study, 7 out of the 29 (24%) of the children in the PT group were

either currently receiving early intervention services or had received early intervention services in the past. Out of these children, none were identified as having a language impairment through standardized testing when they participated in the current study. In other words, although almost a quarter of the children in the PT group had received speech and/or language intervention, none of these children presented with a language impairment (as measured by standardized assessment) when they participated in this current study. It is possible that differences in access to and utilization of early intervention services contributed to the lack of group differences seen on the standardized assessment measures. If these children had not received early intervention services, it could be hypothesized that their deficits may have been greater. However, the present study was not designed to test this hypothesis. Further, the Crosbie et al. (2010) and Mahurin Smith (2014) studies did not disclose rates of intervention utilization, so it is difficult to draw direct comparisons. Most studies of language outcomes of children born preterm do not report rates of utilization of early intervention services. However, in Foster-Cohen's (2010) large sample of preschoolers born preterm, 20.7% of the children who showed no delays on standardized assessments of language as preschoolers had received early intervention.

Considering the results of the present study as a whole, this study offers insight into the language functioning of children born preterm. Many variables related to the neurodevelopmental outcomes of these children including language, attention, and non-verbal intelligence were assessed. Out of all of these aspects of development, the children in the preterm group had significantly poorer outcomes on measures of conversational language skills when assessed through language sample analysis. This indicates that in children born preterm, language difficulties may exist even when these children appear to be developing typically when assessed by standardized assessments of language, cognition, and attention.

Additionally, this study further supports the importance of language sample analysis as a tool for measuring the expressive language skills of young children. Deficits in conversational language skills may be hard to detect through traditional standardized assessments. Although difficulty with finite verb morphology and sentence complexity are seen as hallmark indicators of language impairment in young children, standardized assessments of language may not adequately assess these skills in comparison to data obtained from language samples (Dunn, Flax, & Sliwinski, 1996).

Clinical Implications

The results of this study provide evidence that children born preterm exhibit poorer language skills than children born full term when language skills are measured via language sample analysis. Even for children born preterm with language deficits not seen on standardized measures of language skills, deficits in conversational semantic and grammatical skills may be exhibited. This finding has important clinical implications for practitioners who work with children born preterm. Clinicians should carefully assess the conversational language skills of the preschool children that they are evaluating and not overly rely on the results of standardized assessment when working with this population. A reliance on standardized assessment measures in the diagnosis of language deficits in this population may result in some children not obtaining services to address their expressive language challenges. Rather, language sample analysis in this population seems imperative.

In order to establish a full picture of the functioning of child's linguistic system, it is important to consider the results of standardized assessment in conjunction with performance on measures obtained from language sample analysis. This study adds to the growing body of research that underscores the clinical utility of using language sample analysis in the assessment

of language development. Indeed, several children in this study who showed language deficits on the language sample analysis measures scored within the average range when language was assessed through standardized assessment. Language sample analysis is a valid and reliable assessment of language and is able to detect language deficits that are unable to be detected through standardized assessment (Dunn, Flax, Sliwinski, & Aram, 1996) and remains an invaluable tool.

A recent study investigating the use of language sample analysis in clinical settings showed that speech-language pathologists do not use language sample analysis consistently, and frequently cite barriers such as time constraints as preventing them from using language sample analysis more frequently (Pavelko, Owens, Ireland, & Hahs-Vaughn, 2016). Although language sample analysis can be a lengthy process, its clinical benefit and accuracy in identifying children with language difficulties cannot be overlooked. Researchers should work to develop tools that aid speech-language pathologists to collect, transcribe, and analyze language samples in an efficient manner. Particularly, accurate voice-to-text software coupled with sample analysis software would greatly increase the efficiency of language sample transcription; however, this technology is not yet available for use in the clinical setting. Until these tools can be developed, speech-language pathologists who hope to use language sample analysis more frequently can take several steps to make this process more feasible in the clinical setting. Although this study utilized relatively long samples (100 complete and intelligible utterances), some evidence supports the relative stability of shorter samples (Heilmann, Nockerts, & Miller, 2010). Additionally, age-based normative data are available for many measures that can be obtained from a language sample (Goffman & Leonard, 2000; Rice et al., 2010) and the use of these norms could aid clinical decision-making.

Although this study provides further evidence to support the use of language sample analysis, standardized assessment certainly has a place in the assessment of language. Standardized assessment offers clinicians and researchers a reliable means for assessing receptive language, which is not assessed through language sample analysis. Although receptive language was not within the scope of this study, 5 out of the 29 children in the PT group had Receptive Language Index scores that were below a standard score of 85 (i.e., one standard deviation below the mean). Out of these children, 2 did not have expressive language deficits as determined by either language sample analysis measures or the Expressive Language Index of the CELF-P2. These children would not have been identified if language sample analysis had been used in isolation.

The information gained from this study should be taken into consideration by clinicians who work with children born preterm in the NICU follow-up clinic setting. NICU follow-up clinics are commonly used to provide multidisciplinary care to children who were born preterm. A large survey of American NICU follow-up clinics found that the majority of clinics did not follow children past three years of age (Kuppala, Tabangin, Haberman, Steichen, & Yoltan, 2011). This study also found that “speech” assessment services were only available in 67.2% of the clinics, in contrast to neurodevelopmental outcomes assessments provided in 94.4% of the clinics. Although language is typically assessed in neurodevelopmental assessments such as the Bayley Scales of Infant Development, professionals who are not specially trained in language development and disorders may overly rely on standardized assessments of language and are not likely to be trained in language sample analysis techniques. NICU follow-up clinics should consider employing speech-language pathologists more frequently in their follow-up clinics and ensure that the children are followed through the preschool period.

Limitations of the Current Study

Although this study provides important information on the language skills of preschoolers born preterm, the results of this study should be considered in the context of the limitations of the study. One limitation of this study is that due to characteristics of the sample, the results of this study may not be generalizable to other populations of children born preterm. First, the children who participated in this study were primarily monolingual speakers of standard American English. As American society grows more pluralistic, children are increasingly being raised in households where more than one language is spoken or a non-standard dialect of English is spoken. Because the participants in this study came from primarily monolingual households where mainstream dialects of English were spoken, the results may not be generalizable to the greater population of children born preterm.

Another limitation of this study is that it did not investigate receptive language. There is some evidence that children who were born preterm have poorer receptive language abilities than expressive language abilities, especially children with a history of Chronic Lung Disease (Lewis, Singer, Fulton, Short, Klein, & Baley, 2002; Singer, Siegel, Lewis, Hawkins, Yamashita, & Baley, 2001). It is possible that non-linguistic factors such as attention may impact receptive language scores differently than expressive language scores. More research needs to be conducted to fully understand these complex relationships.

Future Directions

Although the results of this study provide evidence that preschoolers who were born preterm lag behind their full term peers when language skill is measured through language sample analysis, several unresolved questions about language development in this population remain. First, this study found differing outcomes than two previous studies comparing results of

standardized assessment to language sample analysis in children born preterm. These studies found that children born preterm performed more poorly when language skill was measured through standardized assessment than language sample analysis. It is currently unknown if this difference was due to sample characteristics such as gestational age, maternal education or other contributing factors. This problem could be resolved by following this current group of children into the school-aged years to determine if the same pattern is exhibited by this cohort of children when they are older.

In the present study, there was a high degree of variability seen in the group of children born preterm. Some of the children had significant language deficits while others had above average language skills. Many questions remain about which factors, including medical, genetic, and environmental, have the greatest impact on the language outcomes and trajectories of children born. Future studies of language development in children born preterm should work to identify both how specific factors contribute to language outcomes and how these factors interact with each other in increasing or reducing risk. Identifying which factors have the largest impact on subsequent language outcomes would be beneficial for several reasons. First, a better understanding of which factors are most associated with poor outcomes could guide researchers to develop interventions aimed at mitigating these negative factors. Secondly, determining which children born preterm are at the greatest risk for poor language development could result in the early identification of children who would benefit most from early intervention services.

References

- American Academy of Pediatrics (2011). ADHD: Clinical practice guidelines for the diagnosis, evaluation, and treatment of attention-defect/hyperactivity disorder in children and adolescents. *Pediatrics*, 128(5), 1-16.
- American Speech-Language-Hearing Association. (1997). *Guidelines for audiologic screening* [Guidelines]. Available from www.asha.org/policy.
- Aram, D. M., Morris, R., & Hall, N. E. (1993). Clinical and research congruence in identifying children with specific language impairment. *Journal of Speech and Hearing Research*, 36(3), 580-591.
- Archibald, L. M. D., & Joanisse, M. F. (2009). On the sensitivity and specificity of nonword repetition and sentence recall to language and memory impairments in children. *Journal of Speech, Language, and Hearing Research*, 52, 899-914.
- Barlow, S. M., Lee, J., Wang, J., Oder, A., Hall, S., Knox, K., ... Thompson, D. (2014). Frequency modulated orocutaneous stimulation promotes non-nutritive suck development in preterm infants with respiratory distress syndrome or chronic lung disease. *Journal of Perinatology*, 34, 136-142.
- Battle, D. (2002). *Communication disorders in multicultural populations (3rd ed.)*. Boston, MA: Butterworth-Heinneman.
- Beck, I. L., McKeown, M. G., & Kucan, L. (2013). *Bringing words to life, second edition: Robust vocabulary instruction*. New York, NY: The Guilford Press.
- Bhutta, A. T., Cleves, M. A., Casey, P. H., Cradock, M. M., & Anand, K. J. S. (2002). Cognitive and behavioral outcomes of school-aged children who were born preterm. *JAMA*, 288(6), 728-737.

- Bos, A. F., Van Braeckel, K., Hitzert, M. M., Tanis, J. C., & Roze, E. (2013). Development of fine motor skills in preterm infants. *Developmental Medicine & Child Neurology*, 55(s4), 1-4.
- Brown, R. (1973). *A first language: The early stages*. Cambridge, MA: Harvard University Press.
- Butcher, P. R., Kalverboer, A. F., Geuze, R. H., & Stremmelaar, E. F. (2002). A longitudinal study of the development of shift of gaze to a peripheral stimulus in preterm infants with transient periventricular echogenicity. *Journal of Experimental Child Psychology*, 82(2), 116-140.
- Carrow-Woolfolk, E. (2011). *Oral and written language scales (2nd ed.) foundations of language assessment*. Torrance, CA: Western Psychological Services.
- Carvale, B., Tozzi, C., Albino, G., & Vicari, S. (2005). Cognitive development in low risk preterm infants at 3-4 years of life. *Archives of Disease in Childhood – Fetal & Neonatal Edition*, 90, F474-F479.
- Casiro, O., Moddemann, D., Stanwick, R., Panikkar-Thiessen, V., Cowan, H., & Cheang, M. (1990). Language development of very low birth weight infants and fullterm controls at 12 months of age. *Early Human Development*, 24(1), 65-77.
- Cole, K., Mills, P., & Dale, P. (1989). Examination of test-retest and split-half reliability for measures derived from language samples of young handicapped children. *Language, Speech, and Hearing Services in Schools*, 20, 245-258.
- Centers for Disease Control and Prevention (2015). Births: Final data for 2013. *National Vital Statics Reports*, 64(1), 1-65.
- Cohen, J. (1977). *Statistical power analysis for the behavioral sciences*. Routledge.

- Condouris, K., Meyer, E., & Tager-Flusberg, H. (2003). The relationship between standardized measures of language and measure of spontaneous speech in children with autism. *American Journal of Speech-Language Pathology*, 12, 349-358.
- Conners (2015). Conners Kiddie Continuous Performance Test 2nd Edition (K-CPT 2) [computer software]. Tonawanda, NY: Multi-Health Systems Inc.
- Conners C. K., & Staff M. H. (2001) Conners Kiddie Continuous Performance Test (K-CPT): Computer Program for Windows Technical Guide and Software Manual. Toronto, ON: Multi-Health Systems, Inc.
- Conti-Ramsden, G., Botting, N., & Faragher, B. (2001). Psycholinguistic markers for specific language impairment (SLI). *The Journal of Child Psychology and Psychiatry*, 42(6), 741-748.
- Costanza-Smith, A. (2010). The clinical utility of language samples. *SIG 1 Perspectives on Language Learning and Education*, 17, 9-15.
- Crosbie, S., Holm, A., Wandschneider, S. & Hemsley, G. (2011). Narrative skills of children born preterm. *International Journal of Language & Communication Disorders*, 46(1) 83-94.
- Curry, A. E., Pfeiffer, M. R., Slopen, M. E., & McVeigh, K. H. (2012). Rates of early intervention referral and significant developmental delay, by birthweight and gestational age. *Maternal and Child Health Journal*, 16(5). 989-996.
- De Lamo White, C., & Jin, L. (2011). Evaluation of speech and language assessment approaches with bilingual children. *International Journal of Language and Communicator Disorders*, 46(6), 613-627.

de Vries, L. S., Eken, P., Groenendaal, F., van Haastert, I. C., & Meiners, L. C. (1993).

Correlation between the degree of periventricular leukomalacia diagnosed using cranial ultrasound and MRI later in infancy in children with cerebral palsy. *Neuropediatrics*, 24(5), 263-268.

Demuth, K. (1984). *Aspects of Sesotho language acquisition*, Indiana University Linguistics Club: Bloomington.

Dunn, M., Flax, J., Silwinski, M., & Aram, D. (1996) The use of spontaneous language measures as criteria for identifying children with specific language impairment: An attempt to reconcile clinical and research incongruence. *Journal of Speech and Hearing Research*, 39(3), 643-654.

Ebert, K. D., & Scott, C. M. (2014). Relationships between narrative language samples and norm-referenced test scores in language assessments of school-age children. *Language, Speech, and Hearing Services in Schools*, 45, 337-350.

Eisenberg, S. L., Ferkso, T. M., & Lundgren, C. (2001). The use of MLU for identifying language impairment in preschool children: A review. *The American Journal of Speech-Language Pathology*, 10, 323-342.

Eisenberg, S. L., & Guo, Y. (2015). Sample size for measuring grammaticality in preschool children from picture-elicited language samples. *Language, Speech, and Hearing Services in Schools*, 46, 81-93.

Fanaroff, A. A., Stoll, B. J., Wright, L. L., Carlo, W. A., Ehrenkranz, R. A., Stark, A. R., ... Poole, K. (2007). Trends in neonatal morbidity and mortality for very low birthweight infants. *American Journal of Obstetrics & Gynecology*, 196(2), 147.e1-147.e8.

- Fenson, L., Dale, P., Reznick, S., Thal, D., Bates, E., Hartung, ... Reilly, J. (1993). *The MacArthur Communicative Development Inventories*. San Diego, CA: Singular Publishing Group.
- Foster-Cohen, S. H., Friesen, M. D., Champion, P. R., & Woodward, L. J. (2010). High prevalence/low severity delay in preschool children born very preterm. *Journal of Developmental & Behavioral Pediatrics*, 31(8), 658-667.
- Friberg, J. C. (2010). Considerations for test selection: How do validity and reliability impact diagnostic decision? *Child Language Teaching and Therapy*, 26(1), 77-92.
- Gavin, W. J., & Giles, L. (1996). Sample size effects on temporal reliability of language sample measures of preschool children. *Journal of Speech, Language, and Hearing Research*, 39, 1258-1262.
- Goffman, L., & Leonard, J. (2000). Growth of language skills in preschool children with specific language impairment: Implications for assessment and intervention. *American Journal of Speech-Language Pathology*, 9, 151-161.
- Grunau, R. E., Whitfield, M. F., & Davis, C. (2002). Pattern of learning disabilities in children with extremely low birth weight and broadly average intelligence. *Archives of Pediatric & Adolescent Medicine*, 156(6), 615-620.
- Guarini, A., Marini, A., Savini, S., Alessandrini, R., Faldella, G., Sansavini, A. (2016). Linguistic features in children born very preterm at preschool age. *Developmental Medicine & Child Neurology*, 58(9), 949-956.
- Hack, M., Klein, N. K., & Taylor, H. G. (1995). Long-term developmental outcomes of low birth weight infants. *The Future of Children*, 5(1), 176-196.
- Hart, B., & Risely, T.R. (1992). American parenting of language-learning children: Persisting

differences in family-child interactions observed in natural home environments.

Developmental Psychology, 28(6), 1096-1105.

Heilmann, J. (2010). Myths and realities of language sample analysis. *Perspectives on Language Learning and Education*, 17, 4-8.

Heilmann, J., Nockerts, A., & Miller, J. F. (2010). Language sampling: Does the length of the transcript matter? *Language, Speech, and Hearing Services in Schools*, 41, 393-404.

Hughes, D. L., Fey, M. E., & Long, S. H. (1992). Developmental sentence scoring: Still useful after all these years. *Topics in Language Disorders*, 12(2), 1-12.

Kaufman, A. S., & Kaufman, N. L. (2004). *Kaufman brief intelligence test, second edition*. San Antonio, TX: The Psychological Corporation.

Kerr-Wilson, C. O., Mackay, D. F., Smith, G. C. S., & Pell, J. P. (2012). Meta-analysis of the association between preterm delivery and intelligence. *Journal of Public Health*, 34(2), 209-216.

Kilbride, H. W., Thorstad, K., & Daily, D. K. (2004). Preschool outcome of less than 801-gram preterm infants compared with full-term siblings. *Pediatrics*, 113(4), 742-747.

Klee, T. (1992). Developmental and diagnostic characteristics of quantitative measures of children's language production. *Topics in Language Disorders*, 12(2), 28-41.

Klem, M., Melby-Lervåg, M., Hagtvet, B., Halaas Lyster, S., Gustafsson, J., & Hulme, C. (2015). Sentence repetition is a measure of children's language skills rather than working memory limitations. *Developmental Science*, 18(1), 146-154.

Kuppala, V. S., Tabangin, M., Haberman, B., Steichen, J., & Yolton, K. (2011). Current state of high-risk infant follow-up care in the United States: results of a national survey of academic follow-up programs. *Journal of Perinatology*, 32(4), 293-298.

- Laing, S., & Kamhi, A. (2003). Alternative assessment of language and literacy in culturally and linguistically diverse populations. *Language, Speech, and Hearing Services in Schools*, 34, 44-55.
- Le Normand, M. T., & Cohen, H. (1999). The delayed emergence of lexical morphology in preterm children: the case of verbs. *Journal of Neurolinguistics*, 12(3-4), 235-246.
- Leadholm, B., & Miller, J. (1992). *Language sample analysis: The Wisconsin guide*. Madison: WI: Wisconsin Department of Public Instruction.
- Lee, L. (1974). *Developmental sentence analysis*. Evanston, IL: Northwestern University Press.
- Leonard, L. B. (2014). *Children with specific language impairment*. Cambridge, MA: MIT Press.
- Leonard, L. B., Caselli, M. C., Bortolini, U., McGregor, K. K., & Sabbadini, L. (1992). morphological deficits in children with specific language impairment: The status of features in the underlying grammar. *Language Acquisition*, 2(2), 151-179.
- Leonard, L. B., & Finneran, D. (2003). Grammatical Morpheme Effects on MLU: “The same can be less” revisited. *Journal of Speech, Language, and Hearing Research*, 46, 878-888.
- Lewis, B.A., Singer, L.T., Fulton, S., Salvator, A., Short, E. J., Klein, N., & Baley, J. (2002). Speech and language outcomes of children with bronchopulmonary dysplasia, *Journal of Communication Disorders*, 35(5), 393-406
- Limperopoulos, C., Bassan, H., Sullivan, N.R., Soul, J. S., Roberston, R. L., Moore, M., ... du Plessis, A. J. (2008). Positive screening for autism in ex-preterm infants: Prevalence and risk factors. *Pediatrics*, 121(4), 758-765.
- Loeb, D. F., Imgrund, C. M., Lee, J., & Barlow, S. M. (2017). Preterm neurodevelopmental outcomes following orosensory entrainment intervention. Manuscript in preparation.

- Magill-Evans, J., & Harrison, M. J., Van der Zalm, J., & Holdgrafer, G. (2002). Cognitive and language development of healthy preterm infants at 10 years of age. *Physical and Occupational Therapy in Pediatrics*, 22(1), 41-56.
- Mahone, E. M., & Schneider, H. E. (2012). Assessment of attention in preschoolers. *Neuropsychology Review*, 22(4), 361-383.
- Mahurin Smith, J., DeThorne, L. S., Logan, J. A. R., Channell, R. W., & Petrill, S. A. (2014). Impact of prematurity on language skills at school age. *Journal of Speech, Language, and Hearing Research*, 57, 901-916.
- Malvern, D., & Richards, B. (2012). Measures of lexical richness. *The Encyclopedia of Applied Linguistics*. DOI: 10.1002/9781405198431.wbeal0755
- Mangrin, K. S., Horwood, L. J., & Woodward, L. J. (2017). Cognitive development trajectories of very preterm and typically developing children. *Child Development*, 88(1), 282-298.
- McCauley, R. J., & Swisher, L. (1984). Psychometric reviews of language and articulation tests for preschool children. *Journal of Speech and Hearing Disorders*, 49, 34-42.
- McGowan, J. E., Alderdice, F. A., Holmes, V. A., Johnston, L. (2011). Early childhood development of late-preterm infants: A systematic review. *Pediatrics*, 127(6), 1111-1124.
- Ment, L. R., Allan, W. C., Makuch, R. W., & Vohr, B. (2005). Grade 3 to 4 intraventricular hemorrhage and Bayley scores predict outcome. *Pediatrics*, 116(6), e864-e867.
- Miller, J. (1981). *Assessing language production in children*. Boston: MA, Allyn and Bacon.
- Miller, J., & Chapman, R. (2007). *SALT: Systematic analysis of language transcripts*. Madison, WI: University of Wisconsin, Madison.
- Milligan, D. (2010). Outcomes of children born very preterm in Europe. *Arch dis child Fetal Neonatal Ed.*, 95:f, 234-240.

- Miyata, S., MacWhinney, B., Otomo, K., Sirai, H., Oshima-Takane, Y., Hirakawa, & M. Itoh, K. (2013). Developmental sentence scoring for Japanese. *First Language*, 33(2), 200-216.
- Mukerji, A., Shah, V., & Shah, P. S. (2015). Periventricular/intraventricular hemorrhage and neurodevelopmental outcomes: A meta-analysis. *Pediatrics*, 136(6), 1132-1143.
- O'Brien, M., & Nagle, K. (1987). Parents' speech to toddlers: The effect of play context. *Journal of Child Language*, 14(2), 269-279.
- Orton, J., Spittle, A., Doyle, L., Anderson, P., & Boyd, R. (2009). Do early intervention programmes improve cognitive function and motor outcomes for preterm infants after discharge? A systematic review. *Developmental Medicine & Child Neurology*, 51(11), 851-159.
- Owens, R. E. (2010). *Language disorders: A functional approach to assessment and intervention, 5th Edition*. Boston, MA: Pearson
- Paul, R., & Norbury, C. F. (2012). *Language disorders from infancy through adolescence: Listening, speaking, reading, writing, and communicating*. St. Louis, MO: Elsevier.
- Paranka, M. S., Yoder, B. A., & Brehm, W. (1999). Improved outcome of extremely premature infants in the 1990s. *Military Medicine*, 164.8, 568-571.
- Pavelko, S. L., Owens, R. E., Ireland, M., & Hahs-Vaughn, D. L. (2016). Use of language sample analysis by school-based SLPs: results of a nationwide survey. *Language, Speech, and Hearing Services in Schools*, 47, 246-258.
- Pence Turnbull, K. L., & Justice, L. M. (2011). *Language development from theory to practice*. Boston, MA: Allyn & Bacon.
- Plante, E., & Vance, R. (1994). Selection of preschool language tests. *Language, Speech, and Hearing Services in Schools*, 25, 15-24.

- Polin, R. A., Carlo, W. A., & Committee on Fetus and Newborn (2014). Surfactant replacement therapy for preterm and term neonates with respiratory distress. *Pediatrics*, 133(1), 156-163.
- Redmond, S. M., (2004). Conversational profiles of children with ADHD, SLI and typical development. *Clinical linguistics and phonetics*, 18(2), 107-125.
- Reidy, N., Morgan, A., Thompson, D. K., Inder, T. E., Doyle, L. W., & Anderson, P. J. (2013). Impaired language abilities and white matter abnormalities in children born very preterm and/or very low birth weight. *The Journal of Pediatrics*, 162(4), 719-724.
- Reynolds, C. R., & Kamphaus, R. W. (2015). *Behavior assessment system for children, second edition*. San Antonio, TX: Pearson.
- Rice, M. L., Smolik, F., Perpich, D., Thompson, T., Rytting, N., & Blossom, M. (2010). Mean length of utterance levels in 6-month intervals for children 3 to 9 years with and without language impairments. *Journal of Speech, Language and Hearing Research*, 53, 333-349.
- Rice, M. L., & Wexler, K. (1996). Toward tense as a clinical marker of specific language impairment in English-speaking children. *Journal of Speech, Language, and Hearing Research*, 39, 1239-1257.
- Rice, M. L., & Wexler, K. (2001). *Test of early grammatical impairment*. San Antonio, TX: The Psychological Cooperation.
- Rojas, R., & Iglesias, A. (2009). Making a case for language sample: Assessment and intervention with (Spanish-English) second-language learners. *The ASHA Leader*, 14, 10-11.

- Roth, F. P., Speece, D. L., & Cooper, D. H. (2002). A longitudinal analysis of the connection between oral language and early reading. *The Journal of Educational Research*, 95(5), 259-272.
- Sansavini, A., Guarini, A., Justice, L. M., Savini, S., Broccoli, S., Alessandroni, R., & Faldella, F. (2010). Does preterm birth increase a child's risk for language impairment? *Early Human Development*, 86(12), 765-772.
- Scarborough, H. S. (1990). Index of productive syntax. *Applied Psycholinguistics*, 11(1), 1-22.
- Scheiner, A.P., & Sexton, M.E. (1991). Prediction of developmental outcome using a perinatal risk inventory. *Pediatrics*, 88, 1135-1143.
- Schmidhauser, J., Caflisch, J., Rousson, V., Bucher, H., & Latal, B. (2006). Impaired motor performance and movement quality in very-low-birthweight children at 6 years of age. *Developmental Medicine & Child Neurology*, 48(9), 718-72.
- Singer, L.T., Siegel, A.C., Lewis, B., Hawkins, S., Yamashita, T., & Baley, J. (2001). Preschool language outcomes of children with history of bronchopulmonary dysplasia and very low birth weight. *Journal of Developmental and Behavioral Pediatrics*, 22, 19–26.
- Southwood, F., & Russell, A. F. (2004). Comparison of conversation, freeplay, and story generation as methods of language sample elicitation. *Journal of Speech, Language, and Hearing Research*, 47, 366-376.
- Souto, S. M., Leonard, L. B., & Deevy, P. (2014). Identifying risk for specific language impairment with narrow and global measures of grammar. *Clinical Linguistics and Phonetics*, 28(10), 741-756.

- Tomblin, J. B., Records, N. L., & Zhang, X. (1996). A system for the diagnosis of specific language impairment in kindergarten children. *Journal of Speech, Language, and Hearing Research, 39*, 1284-1294.
- Ukrainetz, T. A., & Bloomquist, C. (2002). The criterion validity of four vocabulary test compared with a language sample. *Child Language Teaching and Therapy, 18*(1), 59-78.
- van de Weijer-Bergsma, E., Wijnroks, L., & Jongmans, M. J. (2008). Attention development in infants and preschool children born preterm: A review. *Infant Behavior and Development, 31*(3), 333-351.
- van Noort-van der Spek, I. L., Franken, M. J. P., & Weisglas-Kuperus, N. (2012). Language functions in preterm-born children: A systematic review and meta-analysis. *Pediatrics, 129*(4), 745-754.
- Wagner, C. R., Nettelbladt, U., Sahlen, B., & Niholm C. (2000). Conversation versus narration in pre-school children with language impairment. *International Journal of Language and Communication Disorders, 35*(1) 83-93.
- Walker, D., Greenwood, C., Hart, B., & Carta, J. (1994). Prediction of school outcomes based on early language production and socioeconomic factors. *Child Development, 65*(2), 606-621.
- Watkins, R. V., Kelly, D. J., Harbers, H. M., & Hollis, W. (1995). Measuring children's lexical diversity: Differentiating typical and impaired language learners. *Journal of Speech and Hearing Research, 38*(6), 1349-1355.
- Wickremasinghe, A.C., Hartman, T.K., Voigt, R.G., Katusic, S.K., Weaver, A.L., Colby, C.E., & Barbaresi, W.J. (2012). Evaluation of the ability of neurobiological, neurodevelopmental and social economic variables to predict cognitive outcome in premature infants. *Child:*

Care, Health, and Development, 38(5), 683-689.

Wiig, E., Secord, W., & Semel, E. (2004). *Clinical evaluation of language fundamentals-Preschool, second edition*. San Antonio: TX: Harcourt Assessment.

Wild, K.T., Betancourt, L.M., Brodsky, N.L., & Hurt, H. (2013). The effect of socioeconomic status on the language outcome of preterm infants at toddler age. *Early Human Development*, 89, 743-745.

Werner, E. & Kresheck, J. D. (1983). *Structured photographic expressive language test-II*. Sandwich, IL: Janelle Publications.

APPENDIX A

Toys used during the collection of the language samples



APPENDIX B

Semantic Analysis: High-frequency words were identified by using the MacArthur-Bates database of words produced commonly by toddlers (Fenson et al., 1991)

baa baa	moose	game
choo choo	mouse	glue
cockadoodledoo	owl	pen
grr	penguin	pencil
meow	pig	playdough
moo	pony	present
ouch	puppy	puzzle
quack quack	rooster	story
uh oh	sheep	toy
vroom	squirrel	apple
woof woof	teddybear	applesauce
yum yum	tiger	banana
alligator	turkey	beans
animal	turtle	bread
ant	wolf	butter
bear	zebra	cake
bee	airplane	candy
bird	bicycle	carrots
bug	boat	cereal
bunny	bus	cheerios
butterfly	car	cheese
cat	firetruck	chicken
chicken	helicopter	chocolate
cow	motorcycle	coffee
deer	sled	coke
dog	stroller	cookie
donkey	tractor	corn
duck	train	cracker
elephant	tricycle	donut
fish	truck	drink
frog	ball	egg
giraffe	balloon	fish
goose	bat	food
hen	block	french fries
horse	book	grapes
kitty	bubbles	green beans
lamb	chalk	gum
lion	crayon	hamburger
monkey	doll	ice

ice cream	coat	boo boo
jello	diaper	penis
jelly	dress	shoulder
juice	gloves	tooth
lollipop	hat	toe
meat	jacket	tongue
melon	jeans	tummy
milk	mittens	vagina
muffin	necklace	basket
noodles	pajamas	blanket
nuts	pants	bottle
orange	scarf	box
pancake	shirt	bowl
peas	shoe	broom
peanut butter	shorts	brush
pickle	slipper	bucket
pizza	sneaker	camera
popcorn	snowsuit	can
popsicle	sweater	clock
potato chip	sock	comb
potato	tights	cup
pretzel	underpants	dish
pudding	zipper	fork
pumpkin	ankle	garbage
raisin	arm	glass
salt	belly button	glasses
sandwich	butt	hammer
sauce	cheek	jar
soda/pop	chin	keys
soup	eye	knife
spaghetti	ear	lamp
strawberry	face	light
toast	finger	medicine
tuna	feet	money
vanilla	hair	mop
vitamins	hand	nail
water	head	napkin
yogurt	knee	paper
beads	leg	penny
belt	lips	picture
bib	mouth	pillow
boots	nose	plant
button	owie	plate

purse	stairs	farm
radio	stove	gas station
scissors	table	home
soap	TV	house
spoon	window	movie
tape	washing machine	outside
telephone	backyard	park
tissue	cloud	party
Kleenex	flag	picnic
toothbrush	flower	playground
towel	garden	school
trash	grass	store
tray	hose	woods
vacuum	ladder	work
walker	lawn mower	yard
watch	moon	zoo
basement	pool	aunt
bathroom	rain	baby
bathtub	rock	babysitter
bed	roof	Babysitter's name
bedroom	sandbox	boy
bench	shovel	brother
chair	sidewalk	child
closet	sky	clown
couch	slide	cowboy
crib	snow	daddy
door	snowman	doctor
drawer	sprinkler	fireman
dryer	star	friend
garage	stick	girl
high chair	stone	grandma
kitchen	street	grandpa
living room	sun	lady
oven	swing	mailman
play pen	tree	man
porch	water	mommy
potty	wind	nurse
rocking chair	beach	Child's own name
refrigerator	camping	people
room	church	person
shower	circus	Pet's name
sink	country	police
sofa	downtown	sister

teacher	cover	pour
uncle	cry	pretend
bath	cut	pull
breakfast	dance	push
bye	draw	put
call	drink	read
dinner	drive	ride
give me five!	drop	rip
get	dry	run
gonna get you!	dump	say
potty	eat	see
hi	fall	shake
hello	feed	share
lunch	find	show
nap	finish	sing
night night	fit	sit
no	fix	skate
patty cake	get	sleep
peekaboo	give	slide
please	go	smile
shh/shush/hush	hate	spill
shopping	have	splash
snack	hear	stand
so big!	help	stay
thank you	hide	stop
this little piggy	hit	sweep
turn around	hold	swim
yes	hug	swing
bite	hurry	take
blow	jump	talk
break	kick	taste
bring	kiss	tear
build	knock	think
bump	lick	throw
buy	like	tickle
carry	listen	touch
catch	look	wait
chase	love	wake
clap	make	walk
clean	open	wash
climb	paint	watch
close	pick	wipe
cook	play	wish

work
write
all gone
asleep
awake
bad
better
big
black
blue
broken
brown
careful
clean
cold
cute
dark
dirty
dry
empty
fast
fine
first
full
gentle
good
green
happy
hard
heavy
high
hot
hungry
hurt
last
little
long
loud
mad
naughty
new
nice
noisy

old
orange
poor
pretty
quiet
red
sad
scared
sick
sleepy
slow
soft
sticky
stuck
thirsty
tiny
tired
wet
white
windy
yellow
yucky
after
before
day
later
morning
night
now
time
today
tomorrow
tonight
yesterday
he
her
hers
him
his
I
it
me
mine

my
myself
our
she
that
their
them
these
they
this
those
us
we
you
your
yourself
how
what
when
where
which
who
why
about
above
around
at
away
back
behind
beside
by
down
for
here
inside/in
into
next to
of
off
on
on top of
out

over
there
to
under
up
with
a
all
a lot
an
another
any
each
every
more
much
not
none

other
same
some
the
too
am
are
be
can
could
did
do
does
don't
gonna
going to
got to
gotta

hafta
is
let
lemme
need
trynta
wanna
was
were
will
would
and
because
but
if
so
then

APPENDIX C

Child History Questionnaire

Child's Name: _____ Child's Date of Birth: _____

Parent's Name: _____ Parent Email: _____

Address: _____
_____ Phone Number: _____

Name of Child's Preschool or Daycare: _____

Pregnancy and Birth Information

1. What was the length of your pregnancy (in weeks)?
_____ weeks

2. Did you have any pregnancy complications?

☐ YES ☐ NO

If yes, please describe:

3. Did your child have any complications following birth?

☐ YES ☐ NO

If yes, please describe:

4. Did your child require care in the Neonatal Intensive Care Unit?

☐ YES ☐ NO

If yes, please describe:

Speech and Language Information

1. Do you have any concerns about your child's speech and language development?

☐ YES ☐ NO

If yes, please describe:

2. Has your child ever received special services for his/her speech and/or language development?

☐ YES ☐ NO

If yes, please describe:

3. What languages does your child speak and understand at home?

	0%	5%	10%	15%	20%	25%	30%	35%	40%	45%	50%	55%	60%	65%	70%	75%	80%	85%	90%	95%	100%
English spoken:	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
English understood:	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Spanish spoken:	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Spanish understood:	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Other language spoken:	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Other lang. understood:	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Please specify the
other language: _____

Hearing and Vision

1. Has your child has his or her hearing tested?

☐ YES ☐ NO

If yes, please describe the results:

2. Has your child had ongoing problems with his/her hearing?

☐ YES ☐ NO

If yes, please describe:

4. Has your child had any ear infections?

☐ YES ☐ NO

If yes, how many? _____

5. Does your child currently have an ear infection?

☐ YES ☐ NO

If yes, please describe how it is being treated and the duration:

6. Does your child have vision problems?

☐ YES ☐ NO

If yes, please describe:

If yes, is your child's vision problem corrected while wearing glasses?

☐ YES ☐ NO

General Health

1. Has your child been diagnosed with any type of medical or behavioral condition (e.g., epilepsy, autism, ADHD, learning disability)?

☐ YES ☐ NO

If yes, please describe:

2. Does your child have any allergies?

☐ YES ☐ NO

If yes, please describe:

3. Has your child received special services or therapies in any area other than speech/language development (e.g., occupational therapy, physical therapy, early intervention)?

☐ YES ☐ NO

If yes, please describe:

Other

1. Is there anything else I need to know about your child?

Family and Demographic Information

1. Number of siblings: _____

2. Current marital status of parents:

☐ Married ☐ Separated ☐ Divorced ☐ Living together

3. Mother's highest level of education:

☐ Some primary school ☐ Some college
☐ Some high school ☐ Associate degree (technical college)
☐ High school diploma ☐ Baccalaureate degree (completed college)
☐ GED ☐ Graduate or professional degree

4. Father's highest level of education:

☐ Some primary school ☐ Some college
☐ Some high school ☐ Associate degree (technical college)
☐ High school diploma ☐ Baccalaureate degree (completed college)
☐ GED ☐ Graduate or professional degree

5. Total household income last year:

☐ Less than 9,999 ☐ Between \$60,000 and \$69,999
☐ Between \$10,000 and \$19,999 ☐ Between \$70,000 and \$79,999
☐ Between \$20,000 and \$29,999 ☐ Between \$80,000 and \$89,999
☐ Between \$30,000 and \$39,999 ☐ Between \$90,000 and \$99,999
☐ Between \$40,000 and \$49,999 ☐ Over \$100,000
☐ Between \$50,000 and \$59,999

6. Child's Race

☐ American Indian/Alaskan Native ☐ Asian ☐ Black/African American
☐ Native Hawaiian/Other Pacific Islander ☐ White

7. Child's Ethnicity

☐ Hispanic/Latino ☐ Not Hispanic/Latino